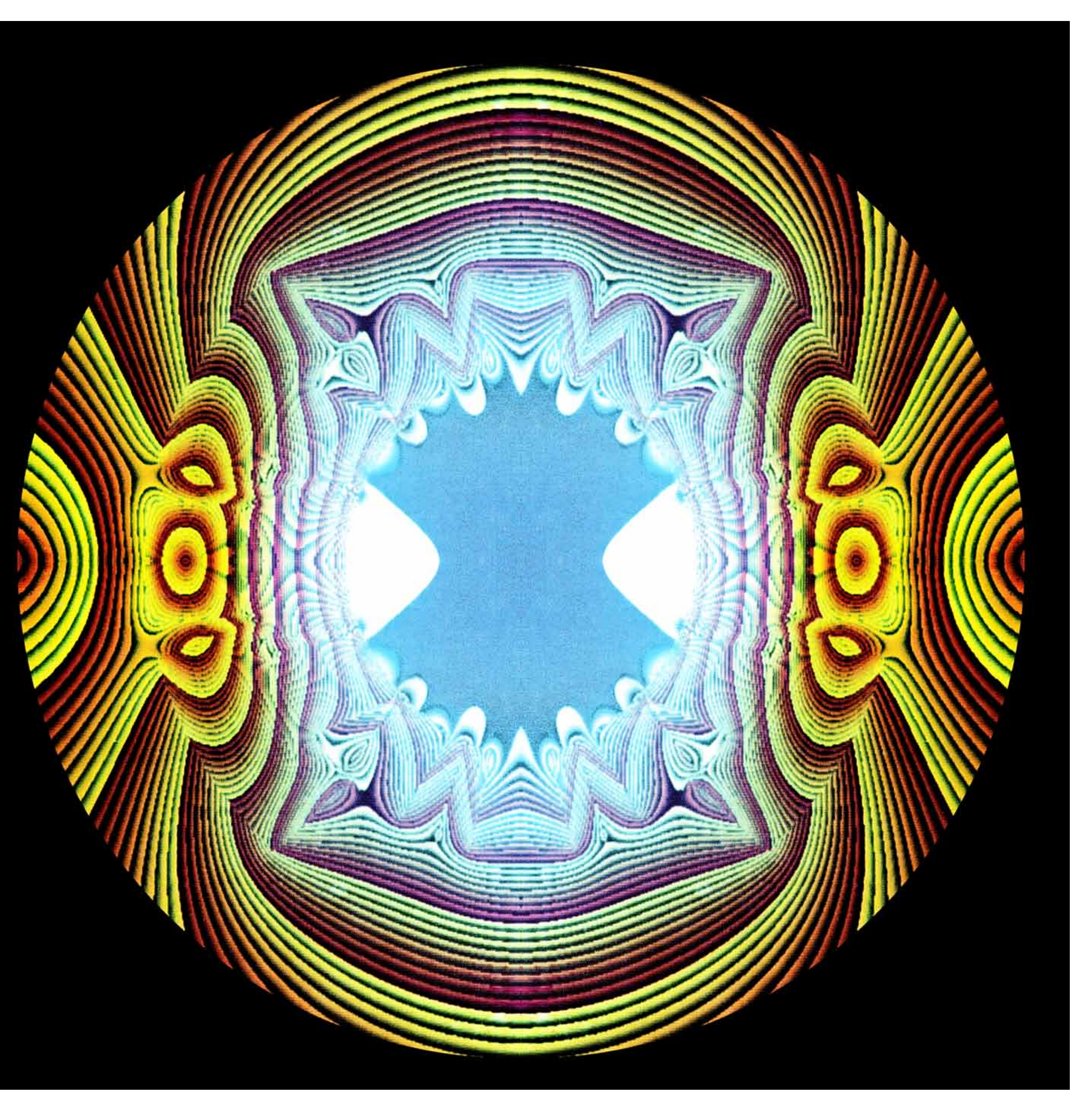
# INVESTIGACION Y CIENCIA

*Edición en español de* SCIENTIFIC AMERICAN



EL SUPERCOLISIONADOR SUPERCONDUCTOR

*Mayo 1986*

450 PTAS.
(IVA INCLUIDO)

Los espacios en gris
corresponden a publicidad
en la edición impresa

*Número 116*
*Mayo 1986*

INVESTIGACION Y CIENCIA *Edición en español de* SCIENTIFIC AMERICAN




EDITA Prensa Científica, S.A.
Calabria, 235-239
08029 Barcelona (ESPAÑA)

**Colaboradores de este número:**

*Asesoramiento y traducción:*

Manuel Puigcerver: *La cola del campo magnético terrestre*; Santiago Torres Martínez: *Genética molecular de la hemofilia*; Ramón Pascual: *El supercolisionador superconductor*; Jordi Ocaña: *Evolución vegetal simulada por ordenador*; José Manuel García de la Mora: *Imágenes mentales y sistema visual*; Gonzalo Ruiz Zapatero: *Almacenaje de grano en Assiros*; Emilio Elizalde: *El efecto Hall cuántico*; Luis Bou: *Juegos de ordenador*; J. Vilardell: *Taller y laboratorio.*

*Ciencia y sociedad:*

Karl von Meyenn

*Libros:*

Luis Bru, Oriol de Bolós, Antonio Domínguez Ortiz y Jesús Mosterín



## PORTADA

La ilustración de la cubierta (de Shlomo Caspi y Michael Helm, del Laboratorio Lawrence Berkeley) es una imagen generada por ordenador del campo magnético que se generaría en el interior de un imán superconductor del proyectado supercolisionador superconductor (SSC) (véase "El supercolisionador superconductor", por J. David Jackson, Maury Tigner y Stanley Wojcicki, en este número). Si se aprueba su construcción, se convertiría en el mayor acelerador de partículas del mundo y permitiría investigar la materia a escalas de longitud y de energía hasta entonces inalcanzables. Los imanes superconductores son un componente clave del diseño de la máquina. Deben curvar dos haces de protones girando en sentidos opuestos en un anillo de unos 85 kilómetros de circunferencia y focalizarlos antes de hacerlos colisionar. La imagen generada por ordenador tiene un color codificado para indicar la intensidad del campo de uno de los imanes de desviación en una sección transversal, tal como podría presentarse a un protón del tubo de vacío por el que circula el haz. Los protones del haz atraviesan la parte central azul de la sección, donde la intensidad del campo es máxima y uniforme.

*Suscripciones:*
Prensa Científica, S. A.
Calabria, 235-239
08029 Barcelona (España)
Teléfono 322 05 51 ext. 33-37

*Condiciones de suscripción:*
España:
Un año (12 números):
4400 pesetas (IVA incluido)
Extranjero:
Un año (12 números): 33 U.S. $
Ejemplar atrasado ordinario:
450 pesetas (IVA incluido)
Ejemplar atrasado extraordinario:
575 pesetas (IVA incluido)

*Distribución para España:*
Distribuciones de Enlace, S. A.
Bruch, 49 - 08009 Barcelona

*Distribución para los restantes países:*
Editorial Labor, S. A.
Calabria, 235-239 - 08029 Barcelona

*Publicidad:*
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137 - 28030 Madrid
Teléfonos 430 84 81 - 437 12 91
Cataluña:
Paulino Muñoz Victoria
Muntaner, 467-469, 3.º, 3.ª
08021 Barcelona
Teléfono 211 92 45

Controlado
por O.J.D.




## PROCEDENCIA DE LAS ILUSTRACIONES

Ilustración de la portada de Shlomo Caspi y Michael Helm

| Página | Fuente |
|---|---|
| 9 | Jorge Garrido, Universidad Católica de Chile |
| 10-17 | Nibaldo C. Inestrosa y Ricardo Génova |
| 12 | Gloria Ruiz, Universidad Católica de Chile |
| 19 | Robert H. Eather, Boston College |
| 20-25 | Ian Worpole |
| 29 | Jon C. Lewis, Facultad de Medicina Bowman Gray, Universidad de Wake Forest |
| 30-35 | Alan D. Iselin |
| 37-40 | Karl von Meyenn |
| 45 | James Kilkelly |
| 46-55 | Gabor Kiss |
| 57-58 | James Kilkelly |
| 59 | Andrew Christie |
| 60-61 | Patricia J. Wynne |
| 62 | Andrew Christie |
| 67 | Stephen Cass |
| 68-73 | Jerome Kuhl |
| 75 | Kenneth A. Wardle, Universidad de Birmingham |
| 76-81 | Tom Prentiss |
| 83-84 | Klaus von Klitzing |
| 85 | Instituto Federal Físico-técnico de Braunschweig |
| 86-93 | Klaus von Klitzing y *Spektrum der Wissenschaft* |
| 95-98 | Edward Bell |
| 101-104 | Michael Goodman |

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa, S.A.
Pedro IV, 160 - 08005 Barcelona
Fotocromos reproducidos por GINSA, S.A.
Gráfica Elzeviriana, S.A.
Nápoles, 249 - Tel. 207 40 11
08013 Barcelona
Printed in Spain - Impreso en España

# Los autores

NIBALDO INESTROSA ("Unión neuromuscular") es profesor de biología celular en la Facultad de Ciencias Biológicas de la Pontificia Universidad Católica de Chile. Nació en 1951, en el sur de Chile, y recibió el grado de doctor en la Universidad Católica en 1979. Durante los dos años siguientes trabajó en el laboratorio de Zach Hall de la Universidad de California en San Francisco. A su regreso de Estados Unidos, se hizo cargo del grupo de neurobiología molecular y fue nombrado jefe del Laboratorio de Neurofisiología, cargo que ocupa en la actualidad. Su campo de investigación ha sido principalmente la biología de la unión neuromuscular, especialmente lo relacionado con la acetilcolinesterasa y la lámina basal sináptica.

EDWARD W. HONES, JR. ("La cola del campo magnético terrestre") es geofísico del Laboratorio Nacional de Los Alamos, Estados Unidos. Cursó estudios de ingeniería mecánica en la Universidad de Duke, por la que se doctoró en 1952. Tras investigar sobre física espacial en la empresa Convair, el Instituto de Análisis de la Defensa y la Universidad de Iowa, ingresó en los Alamos en 1965.

RICHARD M. LAWN y GORDON A. VEHAR ("Genética molecular de la hemofilia") son investigadores de la compañía Genentech. Lawn se graduó en astronomía en el Harvard College, en 1969. Posteriormente cambió su línea de investigación y se doctoró en biología molecular por la Universidad de Colorado en Boulder, en 1977. Lawn realizó trabajos postdoctorales en el Instituto de Tecnología de California, tomando parte en la construcción de la primera genoteca humana. Se incorporó a Genentech en 1980. Vehar se doctoró en 1976 por la Universidad de Cincinnati. Comenzó sus trabajos sobre el factor antihemofílico siendo becario en el departamento de bioquímica de la Universidad de Washington. Desde 1980 está en Genentech.

J. DAVID JACKSON, MAURY TIGNER y STANLEY WOJCICKI ("El supercolisionador superconductor") pertenecen al grupo central de diseño del supercolisionador superconductor (SSC). Jackson, subdirector del mismo, enseña física en la Universidad de California en Berkeley. Se formó en la Universidad de Ontario Occidental y, desde 1949, es doctor en física por el Instituto de Tecnología de Massachusetts. Profesor de las universidades de McGill e Illinois en Urbana-Champaign, se trasladó luego (1967) a Berkeley. Tigner, que dirige el grupo, es catedrático de física de la Universidad de Cornell. Se licenció en ciencias, en 1958, por el Instituto Politécnico Rensselaer, recibiéndose de doctor en física experimental e ingeniería electrónica por la Universidad de Cornell, en 1962. Participa desde hace tiempo en el diseño, construcción y manejo de aceleradores de partículas, tanto norteamericanos como europeos. Wojcicki es jefe del departamento de física de la Universidad de Stanford. Nacido en Polonia, cursó estudios en el Harvard College y se doctoró por Berkeley (1961). Tras investigar en diversos centros, Berkeley incluido, se integró en el claustro docente de Stanford en 1966. También Wojcicki es sudirector del grupo central de diseño.

KARL J. NIKLAS ("Evolución vegetal simulada por ordenador") es profesor asociado de biología de la Universidad de Cornell, donde colabora también en sus departamentos de ecología y sistemática. Cursó el primer ciclo de carrera en el City College de la Universidad de la ciudad de Nueva York; su licenciatura y doctorado, que obtuvo en 1974, son por la de Illinois en Urbana-Champaign. Asistió a la Universidad de Londres en calidad de becario Fullbright-Hayes y, en 1975, fue contratado como conservador del jardín botánico de Nueva York. Niklas ingresó en el claustro docente de Cornell en 1978.

RONALD A. FINKE ("Imagénes mentales y sistema visual") es profesor ayudante de psicología de la Universidad estatal de Nueva York en Stony Brook. Estudió física (1972) y psicología (1974) en la Universidad de Texas en Austin. Se recibió de doctor en psicología por el Instituto de Tecnología de Massachusetts en 1979. En 1981, tras desarrollar trabajos de postdoctorado en Cornell y Stanford, ocupó la plaza de profesor ayudante de psicología de la Universidad de California en Davis. Dos años después Finke se trasladó a Stony Brook, centro en el cual, además de estudiar la imaginería mental, investiga las distorsiones de la memoria visual y la psicología de la creencia en fenómenos paranormales.

GLYNIS JONES, KENNETH WARDLE, PAUL HALSTEAD y DIANA WARDLE ("Almacenaje de grano en Assiros") comparten el trabajo de campo en la expedición arqueológica del yacimiento de Assiros Toumba. Jones se ha especializado en el estudio de restos de plantas; enseña en el departamento de arqueología y prehistoria de la Universidad inglesa de Sheffield. Estudió zoología en la Universidad de Gales, donde obtuvo su licenciatura en 1971. Siguió cursos de postgraduado en arqueología en la Universidad de Cambridge, centro que le otorgó el doctorado en 1983. Kenneth Wardle, director de las excavaciones, da clases en el departamento de historia antigua y arqueología de la Universidad de Birmingham. Se formó en las universidades de Cambridge y de Londres; por esta última se recibió de doctor en arqueología. Halstead, que se especializó en el estudio de restos animales, es profesor en Sheffield, si bien cursó sus estudios en Cambridge, doctorándose en 1984. Diana Wardle, que investiga la cultura material descubierta en el yacimiento, es conservadora del Ministerio Británico de Obras y Construcciones Públicas, denominado en la actualidad Comisión de Monumentos Históricos. Los autores desean manifestar su gratitud a la Escuela Británica de Atenas y al Servicio Arqueológico Griego por la ayuda que les han prestado.

KLAUS VON KLITZING ("El efecto Hall cuántico") dirige, desde 1985, el Instituto Max Planck de Investigación del Estado Sólido, de Stuttgart. Ya en su tesis doctoral, en la que comenzó a trabajar en 1966, abordó la física de los semiconductores; a ella se ha dedicado desde entonces, en especial a las características electrónicas que presentan al someterlos a campos magnéticos intensos. Cursó la carrera en la Universidad de Würzburg, e impartió clases, durante cuatro años, en la Politécnica de Munich. Investigó en Oxford y Grenoble. Sus trabajos en el laboratorio magnético de gran campo de Grenoble, que comparten el Instituto Max Planck y el Centro Nacional Francés de Investigaciones Científicas, llevó al descubrimiento, en 1980, del efecto Hall cuántico; por ello se le otorgó el premio Nobel de física de 1985. Se centran hoy las investigaciones de Klitzing en el campo de las nuevas estructuras de semiconductores.

# Hace...

José M.ª López Piñero

*... doscientos años*

Leonardo Galli publicó el libro titulado *Observaciones de una niña que nació viva, sin cerebro, cerebelo y médula oblongada,* uno de los principales estudios teratológicos realizados en la España de la Ilustración.

Galli nació en Tarragona, en 1751, y recibió su primera formación quirúrgica junto a un cirujano mayor del hospital de dicha ciudad, llamado Salvador Corbella. Estudio más tarde en el Colegio de Cirugía, de Barcelona, donde tuvo entre sus maestros a Francisco Puig y entre sus condiscípulos a Domingo Vidal, otras dos importantes figuras de la cirugía española de estos años con las que mantuvo una duradera relación amistosa y científica. Cirujano militar en los comienzos de su carrera profesional, Galli llegó a ocupar los puestos más elevados del país, entre otros, cirujano de cámara del Rey y director de la Junta Superior Gubernativa de los Colegios de Cirugía. Realizó desde ellos una notable labor, siendo, por ejemplo, uno de los que posibilitó con su informe favorable la célebre expedición dirigida por Francisco Javier de Balmis (1803), que difundió los beneficios de la vacuna. Intervino también en la adopción y propagación de esta última en la misma España, así como en la reglamentación que condujo a la unificación de las profesiones de médico y de cirujano. Falleció en Madrid, en el año 1830.

Las *Observaciones de una niña que nació viva, sin cerebro, cerebelo y médula oblongada* pueden considerarse como el título con el que culminó en nuestro país la etapa preparatoria de la constitución de la teratología como disciplina científica. Tradicionalmente los "monstruos" habían sido objeto de mera curiosidad y su estudio se había limitado, en el mejor de los casos, a su descripción superficial como "caprichos de la naturaleza". El siglo XVIII corresponde a una etapa de transición, en la que se describieron casos con rigurosidad y comenzaron a examinarse las cuestiones biológicas y médicas que planteaban. En España, la centuria se inició en este terreno con la *Observatio rara de corde in monstroso infantulo* (1703), de Martín Martínez, estudio anatomopatológico de un caso de "cor protrusum", que ya hemos comentado en esta misma sección (INVESTIGACIÓN Y CIENCIA, número 98). Se cerró con estas *Observaciones,* en las que Galli se ocupa de una niña que llegó a vivir ocho horas y "cuyo cráneo formaba una masa sólida sin el más mínimo vacío para alojar el cerebro y cerebelo, los cuales faltaban del todo". La observación motiva una serie de reflexiones fisiológicas que expone en una "Memoria sobre los principios de la animalidad", cuya tesis central es que el corazón es el primer motor de la vida orgánica y el órgano más necesario para su mantenimiento. La teratología cristalizó como disciplina científica, superando definitivamente el nivel de curiosidades, cuando pudo ofrecer teorías explicativas basadas en las leyes del desarrollo orgánico que comenzaron a formularse durante la primera mitad del siglo XIX. El libro de Galli pertenece a la fase inmediatamente anterior, en la que los casos teratológicos habían pasado ya de ser considerados como "caprichos de la naturaleza" a ser valorados como "experimentos vivos".

La mentalidad experimentalista de Galli se manifiesta más claramente en otro libro suyo, titulado *Nuevas indagaciones acerca de las fracturas de la rótula y de las enfermedades que con ellas tienen relación, principalmente con la transversal* (1795). El pretexto para publicarlo fue haber asistido su autor a una fractura de dicha clase, que padecía la infanta María Josefa. Está dedicado a Manuel Godoy, que actuó como mecenas de la edición. Corresponde este libro, por tanto, al momento en que el "Príncipe de la Paz" realizó un esfuerzo por volver a la promoción oficial de la ciencia tal como se había practicado en los reinados de Fernando VI y Carlos III. Espléndidamente impreso, es un volumen de más de trescientas páginas, dividido en diez capítulos. El primero es un completísimo estudio de la anatomía de la rodilla. El segundo expone una serie de generalidades acerca de las fracturas. El tercero, el cuarto y el quinto están consagrados, respectivamente, a la fractura transversal de la rótula, a la longitudinal, y a las producidas por contusión, simples y complicadas. El sexto recoge doce observaciones de fractura de la rótula originales de Galli y de otros cirujanos españoles de su tiempo. Los capítulos séptimo y octavo tratan de la formación del callo y de la anquilosis, mientras que el noveno discute las razones de la evolución desfavorable de algunos fracturados de rótula. El capítulo décimo, por último, se dedica a la exposición de algunas enfermedades que tienen relación con la rótula, incluyendo, aparte de los datos y consideraciones del propio Galli, la traducción castellana de unas observaciones realizadas por el cirujano francés Desault.

La exposición del libro de Galli reúne los tres fundamentos habituales en los mejores textos quirúrgicos del siglo XVIII: información amplia y al día, dominio de los datos anatómicos y experiencia clínica sobre la materia. Galli maneja con seguridad la literatura científica de su época, tanto la quirúrgica en sentido estricto, como la relativa a los saberes básicos. Es un experto anatomista práctico que recurre a hechos morfológicos conocidos directamente. Su capítulo acerca de la estructura de la rodilla, por otra parte, constituye, junto a la célebre obra de Gimbernat sobre la hernia crural, uno de los textos más brillantes de anatomía topográfica de la Ilustración española. Su principal fundamento, sin embargo, es su nutrida experiencia clínica que refleja exponiendo un amplio número de casos ilustrativos de los distintos problemas analizados.

La parte más creadora y original de la labor de Galli es que aspiró a explicar las causas y el mecanismo de producción de las fracturas rotulianas mediante la investigación experimental. Realizó, en efecto, series de experimentos, que expone cuidadosamente, consagrados a aclarar tres problemas principales: la fractura transversal de la rótula, la longitudinal y la sección de la porción carnosa del cuádriceps femoral. Véase, como ejemplo, el comentario que dedica a los resultados obtenidos con los experimentos en torno al primero: "Aunque los fenómenos de la fractura transversal de la rótula no sean absolutamente los mismos en el cuerpo vivo que en el cadáver, por carecer en éste los músculos de la fuerza viva, y ser los huesos muertos más frágiles o quebradizos, sin embargo, concurren a producir estos fenómenos otras varias causas, comunes al cuerpo vivo y al muerto, que hacen

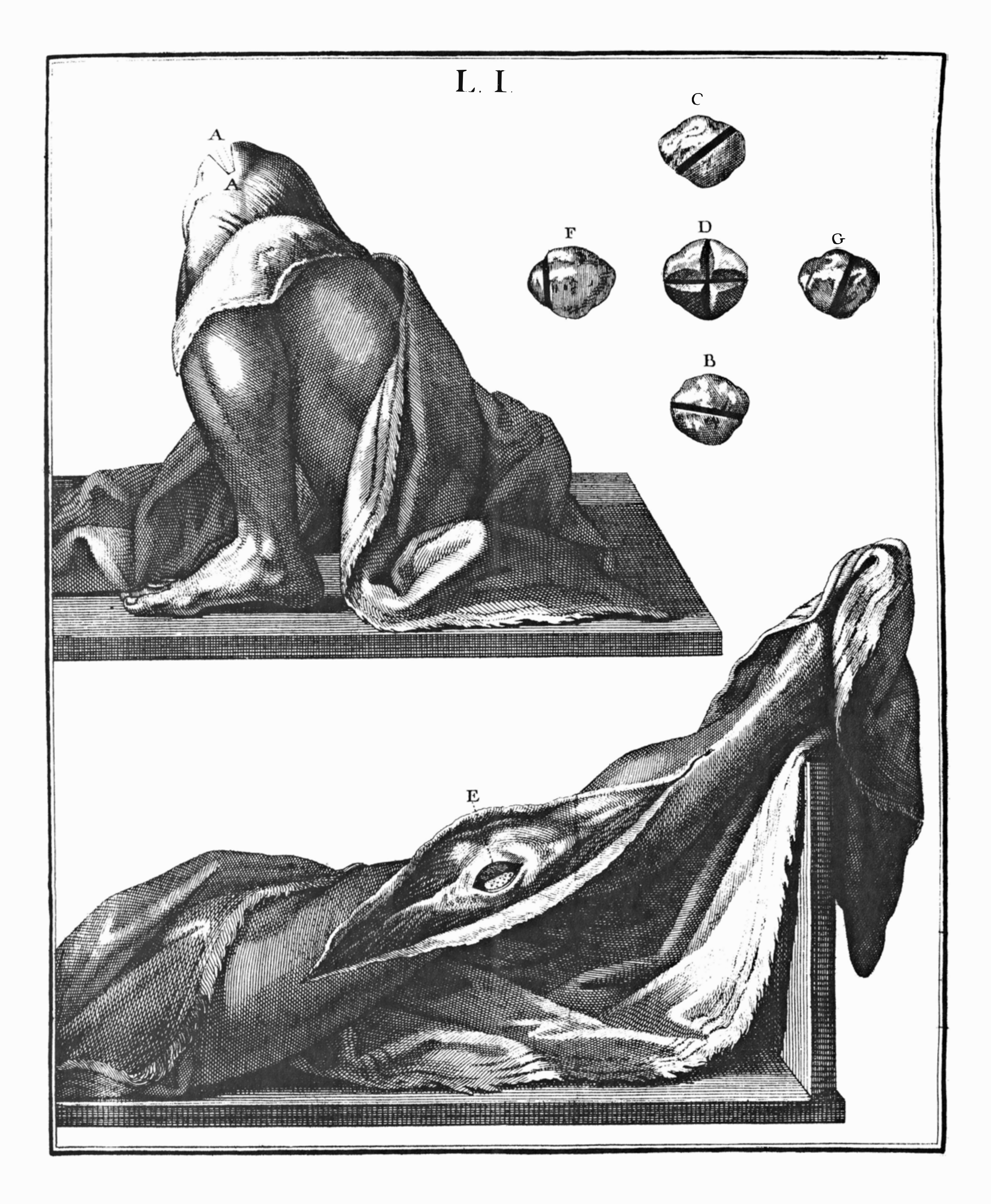



*Lámina de Leonardo Galli relativa a sus investigaciones experimentales, en el cadáver, de las lesiones traumáticas de la rótula: "**AA.** Abertura de los bordes anteriores y posteriores de las piezas de la Rótula fracturada transversalmente, hecha la flexión entera de la pierna con el muslo. **B.** Fractura transverso-oblicua de la Rótula. **C.** Fractura oblicua. **D.** Fractura crucial. **E.** Fractura longitudinal, cuya abertura se manifiesta con la elevación de la pierna. **F.** Fractura longitudinal en su cuarta parte externa. **G.** Fractura longitudinal-oblicua, u oblicuo-longitudinal."*

que estos corolarios sean aplicables a uno y a otro [...]. Como la rótula tiene en algún modo la figura de una castaña y está [...] en el centro de la articulación, por la parte que forma arco o ángulo más o menos agudo en la flexión de la pierna con el músculo, en los movimientos regulares, o cuando el muslo es llevado hacia la pierna en todos los movimientos que el punto de apoyo se efectúa en el pie o en alguna de sus partes, como cuando nos arrodillamos, etc., es forzoso que la rótula se abra en los términos referidos, respecto a que está atada principalmente por la parte anterior de su borde superior, bordes y ángulo inferior. Este mismo mecanismo se evidencia en todo cuerpo que quiere abrírsele por dos pujanzas aplicadas a los lados y que obren en sentido o dirección opuesta, como cuando con las manos se abre una granada, albaricoque, abridor, etc.; y en estos cuerpos no media la circunstancia de estar situados encima de otro que presenta arco o ángulo, como le sucede a la rótula con la articulación del fémur y tibia, que siempre se fractura en el estado anguloso, o de flexión, principalmente en los casos cuya causa no es otra que la contracción violenta del cuádriceps femoral, que remata en la pierna. De todo lo referido, y porque la rótula guarda tal situación en la polea del fémur que a la menor flexión de la pierna presenta la articulación más o menos ángulo, resulta que la rótula transversalmente dividida, tiene precisión de abrirse por los bordes anteriores, a proporción que se está ejecutando la flexión. Mas como en este mecanismo la cara posterior de la rótula no se aparta de su eje o situación, y el cuádriceps femoral queda más o menos relajado, por no efectuarse sino muy poca flexión, se sigue por precisión que el tendón y ligamento que la abrazan superior, inferior y anteriormente, no permiten a la rótula sino el que se la abran los bordes anteriores, en tanto que los posteriores se tocan, en el caso que se ha dicho. [...], me parecieron consecuentes los resultados mencionados, con relación a la situación prominente que tiene la rótula en los extremos grandes y altos hacia adelante, formados por la extremidad inferior del fémur y superior de la tibia, en el concurso de esta articulación. No es menos poderosa la razón de retracción precisa por falta de equilibrio y apoyo, que en el instante de romperse la rótula a la tuberosidad de la tibia [...]. Por lo cual, no es de extrañar que practicadas las flexiones, no se toquen los bordes anteriores en la extensión de la pierna, cuyo efecto debe ser más cierto en los vivientes; como igualmente lo es el de la proporción de tres a uno de los bordes anteriores a los posteriores en aquellos individuos que yo he reconocido, en quienes la rótula ha quedado sin reunirse. Siguiendo la rótula los movimientos del cuádriceps femoral, y sujeta en su natural sitio por las potencias referidas [...], no me parece extraño que se rompa transversalmente dicho hueso, aun en aquel estado de flexión de la pierna, en el cual nos parece que el cuádriceps femoral está relajado, como efectivamente lo está en la simple flexión sin apoyo de la pierna y pie. Mas si en esta misma flexión nos apoyamos sobre la pierna y pie para resistir el peso del cuerpo, notaremos que en el acto de la resistencia entra en contracción el cuádriceps femoral y que la rótula se halla fuertemente retenida por el tendón de dicho músculo y por el ligamento tendinoso. De modo que la rótula tendrá que sufrir la extraordinaria opresión y fuerza de ambas potencias opuestas y todo el peso del cuerpo. Si este hueso, pues, se halla en algún paraje en estado de no poder tolerar y resistir estas dos fuerzas opuestas, se romperá precisamente por aquel punto, partiéndose en dos pedazos, de los cuales el uno se lo llevará hacia arriba el tendón del cuádriceps y el otro hacia abajo el ligamento tendinoso, con el mecanismo referido [...]. El cuádriceps femoral con su tendón, rótula y ligamento tendinoso, puesta la pierna en extensión, describe una curva hacia la rótula, cuya convexidad mira al cóndilo interno del fémur; lo que hace que sean más frecuentes las fracturas transverso-oblicuas de este hueso. Pero esta sabia colocación de dichas partes tiene la ventaja de que puede la pierna ejercer mayores flexiones [...]. Las fracturas transverso-oblicuas de la rótula presentan, a poca diferencia, los mismos fenómenos y dificultades que las transversales..."

Este aspecto de la obra de Galli la sitúa en la vanguardia científica de la cirugía de la Ilustración, concretamente dentro de la tendencia encabezada por el británico John Hunter, que intentaba fundamentar el saber y la práctica quirúrgicos, no sólo sobre los conocimientos anatómicos y los datos clínicos, sino en los resultados de la investigación experimental. Por ello, su aportación no fue apreciada únicamente en España, como lo demuestra, entre otras cosas, la traducción italiana de las *Nuevas indagaciones* que se publicó tres años después de la edición original.

# Unión neuromuscular

*Nervio y músculo se hayan implicados en el control de la placa motora. Sin embargo, una estructura apenas conocida, la lámina basal, desempeña un papel clave en el mantenimiento y diferenciación de las membranas*

Nibaldo C. Inestrosa

Las células nerviosas, o neuronas, amén de comunicarse entre sí, lo hacen también con las células musculares. A mediados del siglo pasado se había observado ya que, cuando las fibras de un nervio motor se acercaban a las fibras musculares estriadas, los axones perdían su vaina de mielina y, cuando entraban en contacto con la superficie muscular, lo hacían en una zona que era particularmente rica en núcleos. Luego de este primer contacto, el nervio se ramificaba dando origen a una arborización terminal, que variaba de músculo en músculo y también de especie en especie. Posteriormente, con los trabajos realizados a comienzos de este siglo por Jorge Francisco Tello y Santiago Ramón y Cajal sobre la génesis de las terminaciones nerviosas motoras, se pudo establecer que la diferenciación de estas estructuras ocurría tempranamente en el desarrollo.

La unión neuromuscular entre una motoneurona y una célula muscular difiere de una sinapsis interneuronal en varios aspectos. Por su origen: las células nerviosas y musculares derivan de diferentes capas germinales embrionarias; por su distribución: las membranas muscular y nerviosa están separadas por un espacio de 500 angstrom, ocupado por una matriz extracelular amorfa; y por su constitución: el terminal nervioso está cubierto por un tipo de células gliales que no se encuentra en el sistema nervioso central. Sin embargo, cada principio que gobierna la función de las sinapsis interneuronales químicas se ha definido también en la unión neuromuscular; por cuya razón, la descripción de la formación y regulación de la unión neuromuscular puede aplicarse a las conexiones que ocurren en el sistema nervioso central. Más aún, muchas preguntas relacionadas con el desarrollo y regulación de los contactos entre nervio y músculo son repeticiones de otras relativas a problemas más generales de la biología, como lo es el reconocimiento e interacción específica entre diversos tipos celulares.

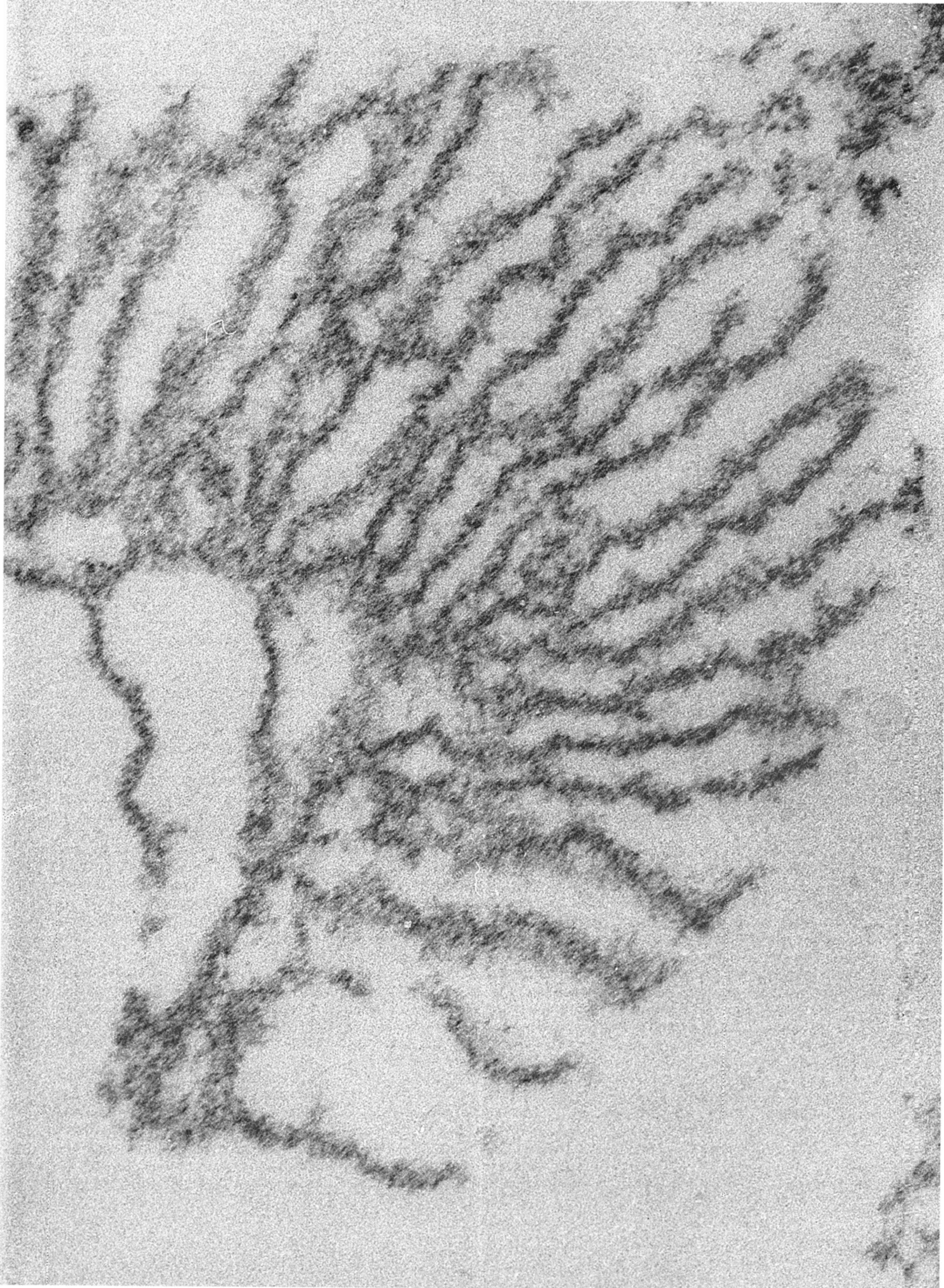
La sinapsis de la unión neuromuscular implica una compleja y precisa interacción entre las células que participan en ella. Estas interacciones se manifiestan a través de la especialización del terminal axónico presináptico, de la membrana postsináptica y del material que ocupa el espacio entre ellos. La mayoría de las uniones neuromusculares, a pesar de sus variaciones en geometría, son similares en ultraestructura. En general, los terminales nerviosos están cubiertos por procesos de las células gliales de Schwann y contienen distintos orgánulos, incluyendo mitocondrias y vesículas de 30 a 60 nanómetros de diámetro denominadas vesículas sinápticas. Estas últimas contienen el neurotransmisor acetilcolina (ACh), que constituye el mensaje normalmente transmitido desde el nervio hasta el músculo. Muchas de las vesículas se localizan cerca de zonas densas de la membrana presináptica, en el lado opuesto a la fibra muscular. Nos referimos a las zonas activas, en las cuales las vesículas se fusionan con la membrana plasmática para liberar el neurotransmisor. Por debajo del terminal nervioso, la membrana de la fibra muscular presenta, a intervalos regulares, ciertas hendiduras: los pliegues sinápticos. Estos se abren precisamente en sitios opuestos donde se encuentran las zonas activas de la membrana presináptica.

La membrana postsináptica posee partículas agregadas que corresponden a proteínas estructurales. La superficie de la fibra muscular está recubierta, de la base al ápice, por la membrana plasmática, el glicocálix, la lámina basal y la lámina reticular. El conjunto de estas diferentes capas se denomina sarcolema. La lámina basal mide de 30 a 40 nanómetros de ancho, presenta un aspecto amorfo y está separada de la membrana plasmática muscular por el glicocálix (10 nanómetros). Por encima de la lámina basal hallamos la lámina reticular, que es una estructura rica en filamentos de colágeno. Ambas láminas integran la membrana basal; sin embargo, sólo la lámina basal junto al glicocálix y la membrana plasmática se extienden a través de la unión neuromuscular y en los pliegues sinápticos. Conviene recordar que, cuando se lesiona el músculo esquelético, las células musculares degeneran y son fagocitadas, pero las vainas de la lámina basal persisten y sirven de esqueleto para la regeneración ordenada de las nuevas miofibrillas.

A diferencia de la membrana muscular sináptica, cuya ultraestructura distínguese de la membrana extrasináptica, las láminas basales de la región sináptica y extrasináptica no poseen características estructurales diversas. Sin embargo, experimentos que se describirán en detalle más adelante muestran que la lámina basal sináptica es una estructura funcionalmente especializada, lo que sugiere una diferencia con las zonas extrasinápticas.

El componente más importante de la membrana postsináptica de la unión neuromuscular, desde el punto de vista funcional y estructural, es el receptor de acetilcolina (AChR): complejo macromolecular que da cuenta tanto del reconocimiento del neurotransmisor como de la translocación de iones a través de la membrana. Al

**1. MEMBRANA BASAL SINAPTICA obtenida del órgano eléctrico de *Discopyge tschudii*. En esta micrografía electrónica tomada por Jorge Garrido, del Laboratorio de Histología de la Universidad Católica de Chile, la membrana basal conserva su citoarquitectura después de ser purificada por centrifugación diferencial y tratamientos con detergentes y alta fuerza iónica. El aumento es de unos 160.000 diámetros.**



| PROPIEDADES | RECEPTORES SINAPTICOS | RECEPTORES EXTRASINAPTICOS |
|---|---|---|
| Densidad normal | $3 \times 10^4$/um$^2$ | $3 \times 10^2$/um$^2$ |
| Recambio metabólico | 11 días | 1 día |
| Tiempo medio de apertura de canales | 1 milisegundo | 4 milisegundos |
| Punto isoeléctrico | 5,1 | 5,3 |
| Determinantes inmunoespecíficos | – | +++ |

**2. DIFERENCIAS entre los receptores de acetilcolina localizados en la unión neuromuscular y en zonas extrasinápticas. La abreviatura *um* es de una unidad de medida que corresponde a $10^{-6}$ metros.**

| REGIONES SINAPTICAS | COMPONENTES COMUNES | REGIONES EXTRASINAPTICAS |
|---|---|---|
| Acetilcolinesterasa | Laminina | Colágeno tipo V |
| Cápsula del cristalino | Colágeno tipo IV | Nuevo tipo de colágeno muscular |
| Colágeno de membrana basal muscular (adsorbido)$^{+}$ | Fibronectina | |
| Colágeno de cápsula del cristalino (adsorbido)$^{+}$ | | |

**3. COMPONENTES MOLECULARES de regiones sinápticas, extrasinápticas y ambas reconocidos por anticuerpos que se unen a la membrana basal de la fibra muscular. Los sueros que llevan un signo (+) son específicos de la unión neuromuscular sólo después de adsorción con tejido conectivo extrasináptico.**

AChR compete interactuar con la ACh liberada por el terminal nervioso y mediar la transmisión de información entre nervio y músculo. Los receptores están concentrados en la membrana postsináptica de la unión neuromuscular, en las crestas de los pliegues sinápticos. El receptor es un oligómero de 250.000 dalton de peso molecular, compuesto por cinco subunidades, de las cuales la más importante es una de 40.000 dalton denominada subunidad alfa. Dicha subunidad contiene el sitio de unión para la ACh y la alfa-hungarototina, una proteína básica que se une en forma casi irreversible al receptor; se obtiene del veneno de serpientes. El AChR es una proteína estructural de la membrana postsináptica, extendiéndose a través de ella de manera que parte de su estructura queda expuesta al exterior y parte al interior de la célula muscular.

La acetilcolinesterasa (AChE) es la macromolécula responsable de destruir la ACh dejando libres a los AChR para una activación posterior. Esta enzima se encuentra en la lámina basal, tanto en el espacio sináptico como en el interior de los pliegues sinápticos. La AChE se solubiliza de la unión neuromuscular mediante tratamiento con la enzima colagenasa, en condiciones tales que el receptor permanece unido a la membrana postsináptica y sus propiedades fisiológicas se mantienen inalteradas. De ello se ocuparon, en 1971, Zach Hall y Regis Kelly, de la Universidad de Harvard, quienes plantearon la posibilidad de que la AChE estuviera inmersa en una matriz de colágeno.

Al igual que los receptores de acetilcolina, la AChE de los órganos eléctricos de ciertos peces se ha purificado y caracterizado. Especial mención merecen los estudios de Jean Massoulie y colaboradores, en la Escuela Normal Superior de París, quienes han establecido la existencia de formas globulares y asimétricas de la AChE. Las formas globulares se clasifican en $G_1$, $G_2$ y $G_4$, donde el número representa la cantidad de subunidades que contiene cada forma. Cada subunidad presenta un peso molecular de 70.000 dalton.

Las formas asimétricas, por su parte, se denominan $A_4$, $A_8$ y $A_{12}$. El carácter asimétrico lo determina un tallo de 500 angstrom de longitud, con una estructura similar al colágeno, tal como lo han demostrado Terry Rosenberry, de la Universidad de Columbia, Palmer Taylor, de la Universidad de California, e Israel Silman, del Instituto Weizman de Israel. La AChE no se ha purificado aún de músculos estriados de vertebrados. Sin embargo, tal como lo demostró Zach Hall en 1973, existe en ellos una multiplicidad de formas de AChE, de las cuales la forma asimétrica $A_{12}$ que posee un coeficiente de sedimentación de 16$S$ se localiza de manera preferencial en la unión neuromuscular. (El coeficiente de sedimentación es un parámetro de caracterización molecular que considera el peso molecular y la simetría de una macromolécula. La unidad de medición se designa como $S$ y equivale a $10^{-13}$ Svedberg.)

En el desarrollo de las células musculares estriadas existe inicialmente una fase de proliferación en la cual las células mononucleadas (mioblastos) comienzan a dividirse. Posteriormente, en la fase de diferenciación, los mioblastos tienden a fusionarse y formar grandes células multinucleadas, denominadas miotubos, que no se vuelven a dividir.

En la formación de la unión neuromuscular se cubren las siguientes etapas. Primera: sólo existen células no diferenciadas y mioblastos que ocupan áreas llamadas "miotomos", en que posteriormente se desarrollarán los músculos esqueléticos. En los miotomos comienzan a entrar paquetes de fibras nerviosas sin presentar signos de especialización; es más, durante este período los axones pioneros no entran en contacto con los mioblastos. Segunda: los mioblastos se fusionan formando miotubos, las proteínas contráctiles comienzan a sintetizarse y algunas miofibrillas aparecen en el citoplasma. Los axones establecen contacto con los miotubos y el espacio entre nervio y músculo es más estrecho que el observado en el animal adulto (10 a 20 nanómetros). En algunos casos se observa acoplamiento eléctrico.

A lo largo de la tercera etapa, se observan los primeros esbozos de lo que serán las uniones neuromusculares en el animal adulto. Los axones entran en contacto con la superficie miotubular a través de una simple terminación nerviosa, a menudo dirigida justo por encima del núcleo de la célula muscular. El espacio sináptico es ahora de 50 a 100 nanómetros, similar al observado en el animal adulto, y contiene un material amorfo que probablemente representa los primeros signos de la lámina basal sináptica. Los terminales axónicos cubiertos por procesos de las células gliales de Schwann comienzan a acumular vesículas de 50 nanómetros. En la membrana neural que da al miotubo aparecen condensaciones que corresponden a lo que serán más tarde las zonas activas. La célula muscular también adquiere especializaciones en el punto de contacto con el nervio, con engrosamientos de la membrana postsináptica y acumulación de AChR y AChE. Se detecta transmisión sináptica en estos contactos nervio-músculo, de naturaleza química, cuántica (es decir, el neurotransmisor se libera en paquetes discretos) y colinérgica nicotínica.

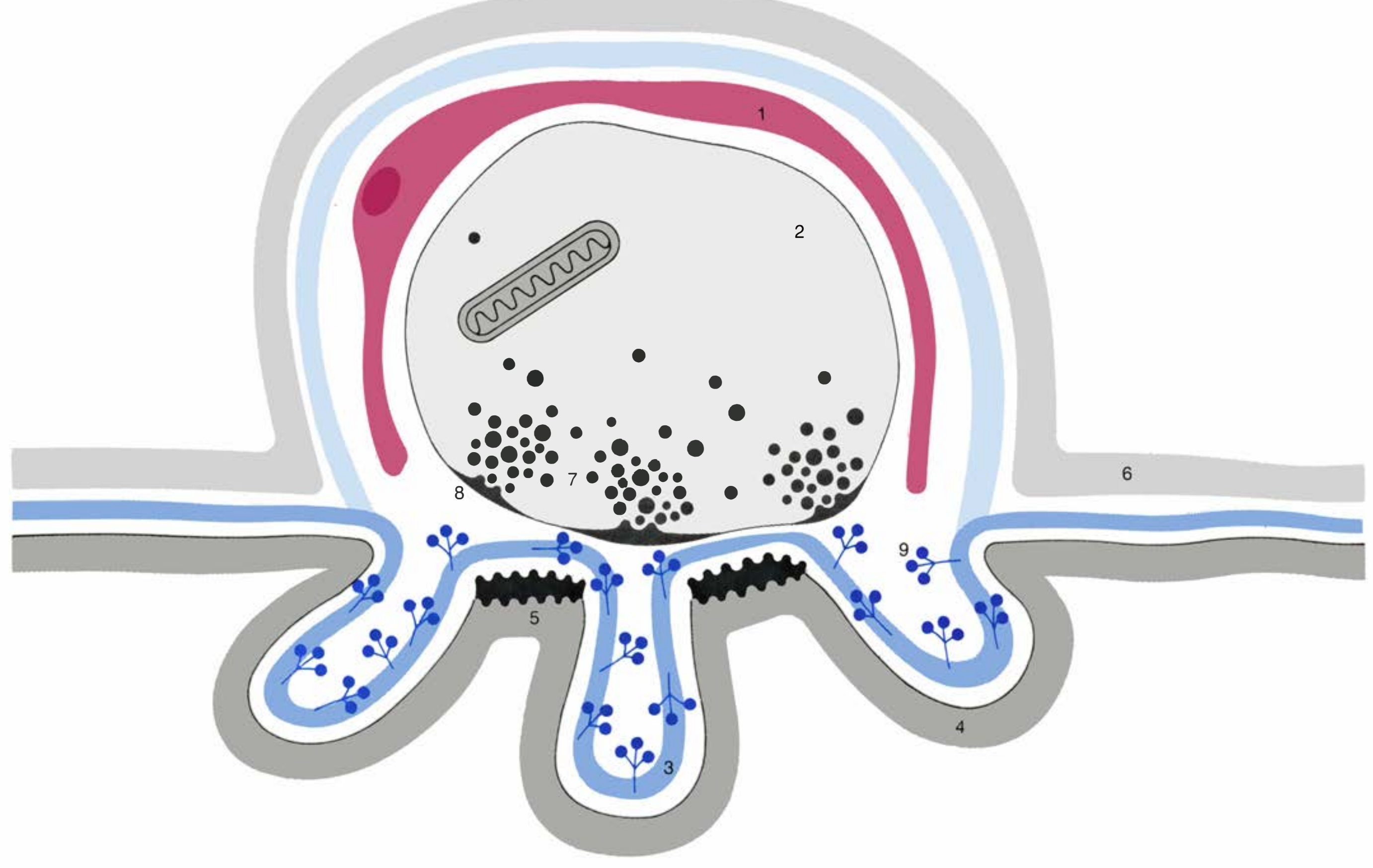
A medida que la unión madura, los terminales nerviosos aumentan de tamaño, presentan numerosas mitocondrias y vesículas sinápticas y las zonas activas continúan su desarrollo: nos hallamos en la cuarta etapa. En el citoplasma de la célula muscular, las mio-

fibrillas se encuentran mejor organizadas y comienza la migración de núcleos desde el centro hacia la periferia del citoplasma. Este movimiento de orgánulos tiene especial significado, ya que define la transición de miotubo a miofibrilla. Por su parte, la lámina basal rodea cada miotubo; AChR y AChE se concentran, abundantemente, en la unión neuromuscular.

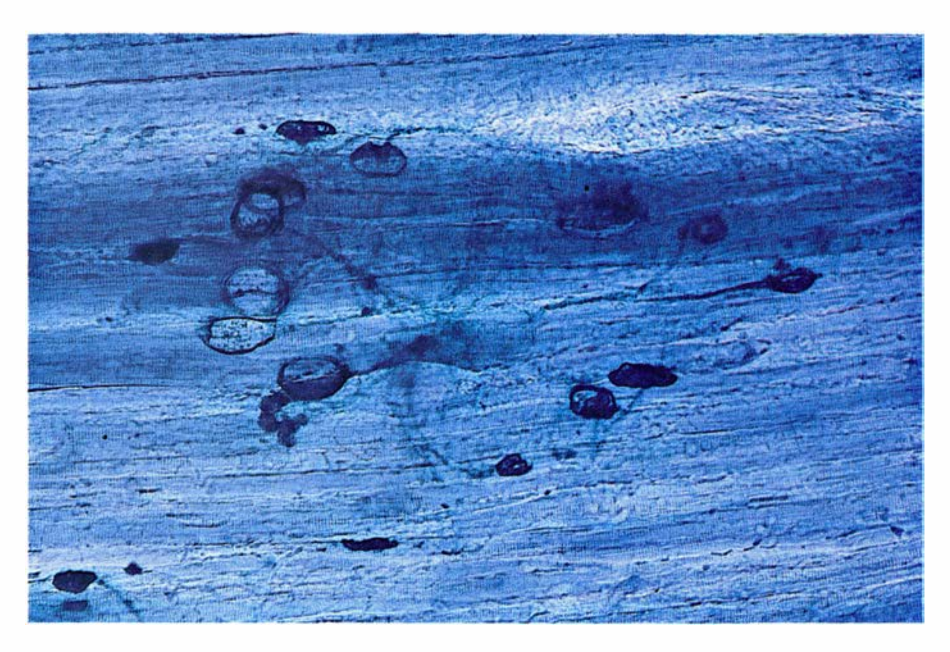
En la etapa siguiente, la quinta, la superficie de la fibra muscular presenta depresiones por debajo del terminal nervioso, las cuales comienzan a formar los pliegues sinápticos. La lámina basal sigue a estos pliegues y la AChE aparece asociada a ella. Por su parte, la membrana de la célula de Schwann se retrae, el terminal crece y numerosas vesículas se concentran y ordenan en las zonas densas de la membrana presináptica, que definitivamente forman las zonas activas.

El estudio de los AChR reviste especial interés. Trátase de una macromolécula que sufre cambios en su concentración y en su localización en la superficie muscular, durante el desarrollo embrionario y la denervación de músculos esqueléticos. Las fibras musculares embrionarias presentan sensibilidad a la ACh a lo largo de toda la superficie, ya que poseen receptores homogéneamente distribuidos en la membrana celular. En etapas tempranas del desarrollo embrionario, el miotubo presenta una distribución homogénea de receptores (160 sitios por $um^2$: um es la abreviatura de una unidad de medida que corresponde a $10^{-6}$ metros). Sin embargo, luego que los axones motores establecen contacto con la superficie muscular, los AChR se acumulan en un sitio determinado, y después de un día de la llegada del nervio, la densidad total de receptores aumenta de 50 a 100 veces. Una vez establecida esta alta densidad de AChR en el lugar de contacto, permanece constante mientras los receptores extrasinápticos desaparecen a una velocidad también constante. El área de la unión neuromuscular representa sólo una pequeña fracción del área total de la fibra muscular (0,1 por ciento), por lo que una importante disminución en el número total de sitios receptores por músculo acompaña la localización del AChR. En efecto, al menos el 95 por ciento del total de receptores se pierde durante este proceso, lo que indica que la formación de la unión neuromuscular ocurre como consecuencia de una severa selección molecular.

En el músculo inervado los AChR están concentrados en la unión neuromuscular. Ahora bien, días después de seccionar un nervio motor, nuevos receptores aparecen en zonas extrasinápticas. Por tanto, una manera de entender la bioquímica de la localización del AChR durante el desarrollo es comparar las propiedades de los receptores de la unión neuromuscular con aquellos que aparecen en zonas extrasinápticas. Con esta idea por delante, Zach Hall y colaboradores, primero en la Universidad de Harvard y luego en la Universidad de California en San Francisco, han estudiado y caracterizado




**4. UNION NEUROMUSCULAR, en representación esquemática. Consta de los siguientes elementos: un terminal nervioso presináptico, que contiene vesículas sinápticas que se ubican en las zonas activas donde se libera la acetilcolina, una célula glial de Schwann, que cubre el terminal nervioso, y un elemento postsináptico, el músculo, que posee una serie de invaginaciones, denominadas pliegues sinápticos. Entre el terminal nervioso y el músculo se encuentra la llamada lámina basal, elemento constituido por proteínas colagenosas, glicoproteínas de alto peso molecular, de las que cabe citar la acetilcolinesterasa y proteoglicanes. La lámina basal presenta pliegues que se continúan con los pliegues sinápticos de la fibra muscular y se advierten donde se libera el neurotransmisor.**



**5. UNION NEUROMUSCULAR DE MUSCULO SOLEO DE RATA, teñida con técnicas histoquímicas específicas para acetilcolinesterasa. Las fibras nerviosas que forman los terminales se visualizan al teñir la preparación con tinción de Giemsa. El aumento es de unos 1.600 diámetros. (Fotografía de Gloria Ruiz.)**

una serie de diferencias entre ambos tipos de receptores. A saber: (*a*) velocidad de degradación; los receptores de la unión neuromuscular se degradan mucho más lentamente (vida media de 11 días) que los receptores extrasinápticos (vida media de 1 día); (*b*) punto isoeléctrico: los receptores extrasinápticos poseen un punto isoeléctrico 0,2 unidades más alto que el del receptor de la unión neuromuscular (5,3 frente a 5,1); (*c*) propiedades inmunológicas: inmunoglobulinas de pacientes con miastenia gravis se interponen a la unión de alfa-bungarotoxina con receptores extrasinápticos, sin alterar la unión de la toxina a los receptores de la unión neuromuscular. Esto implica que los receptores extrasinápticos poseen determinantes específicos que no están presentes en los receptores de la unión neuromuscular.

Se ha investigado también el tiempo medio de abertura de canales iónicos; (*d*): el tiempo medio en que los canales iónicos del AChR permanecen abiertos en respuesta a la ACh es de 1 milisegundo en músculos normales y aumenta a 4,6 milisegundos en músculos denervados; esta diferencia fue descubierta por Bert Sakmann en el Instituto Max Planck de Alemania. En vista de que muchas de las propiedades de los receptores que aparecen con denervación son similares a las observadas en músculos en desarrollo, es posible estudiar en qué momento del período embrionario aparecen los receptores con características de receptores sinápticos. Se sabe que, en la rata, el cambio en la velocidad de degradación del receptor ocurre a los 19 días de gestación, que el cambio en las propiedades inmunológicas ocurre a fines de la primera semana de nacimiento y, finalmente, que el tiempo medio de abertura de canales iónicos se modifica entre la primera y segunda semana después del nacimiento.

En relación a la regulación de los niveles de AChR, Terje Lomo, del Instituto de Neurofisiología de la Universidad de Oslo, ha establecido que la estimulación eléctrica de músculos adultos denervados elimina la presencia de los AChR. Respecto al efecto directo del nervio sobre la síntesis y localización de los AChR en la unión neuromuscular, varios grupos de investigadores han comunicado que extractos neurales contienen factores capaces de afectar el número, la densidad y la agregación de receptores de acetilcolina en músculos en cultivo. Estudios de Miriam Salpeter, de la Universidad de Cornell en Nueva York, han demostrado que la denervación de músculos estriados determina un aumento en la velocidad de degradación de los receptores de la unión neuromuscular. Estos resultados permiten concluir, en términos generales, que el nervio condiciona la síntesis y localización del AChR, al menos de dos formas: mediante factores neurales, en el caso del receptor de la unión neuromuscular, y a través de la actividad eléctrica inducida en el músculo, en el caso del receptor extrasináptico.

Junto a Lea Ziskind-Conhaim y Zach Hall, en la Universidad de California en San Francisco, hemos estudiado recientemente el desarrollo de la AChE en músculos intercostales de embriones de rata. Los resultados indican que la acetilcolinesterasa muscular aparece antes que los AChR. En la rata, ya el día 13 de gestación se detectan formas globulares de la enzima; posteriormente, el día 14, aparece la forma asimétrica $A_{12}$ que posee un coeficiente de sedimentación de 16*S* y es característica de la unión neuromuscular en el animal adulto. Sin embargo, y a pesar de encontrarse esta forma enzimática localizada en el exterior de la célula muscular y ser fisiológicamente activa, su focalización en un sitio específico de la fibra muscular es un proceso muy lento que, si bien comienza entre los días 16 y 17 de gestación, no termina hasta finales del primer mes de desarrollo postnatal.

Hemos estudiado también el desarrollo de la AChE en líneas celulares de músculos esqueléticos en cultivo. Una de estas líneas posee grandes ventajas para el estudio de la expresión y regulación de la AChE 16*S*. En efecto, los mioblastos presentan una actividad mínima de AChE. Sin embargo, una vez que la fusión celular ha comenzado, la actividad aumenta rápidamente llegando hasta 20 veces el valor observado en mioblastos. Concomitante con esto se observa la aparición de la forma 16*S*. Esta forma de la AChE se sintetiza en completa ausencia de neuronas, lo que demuestra que la influencia neural no es fundamental en la síntesis de esta macromolécula.

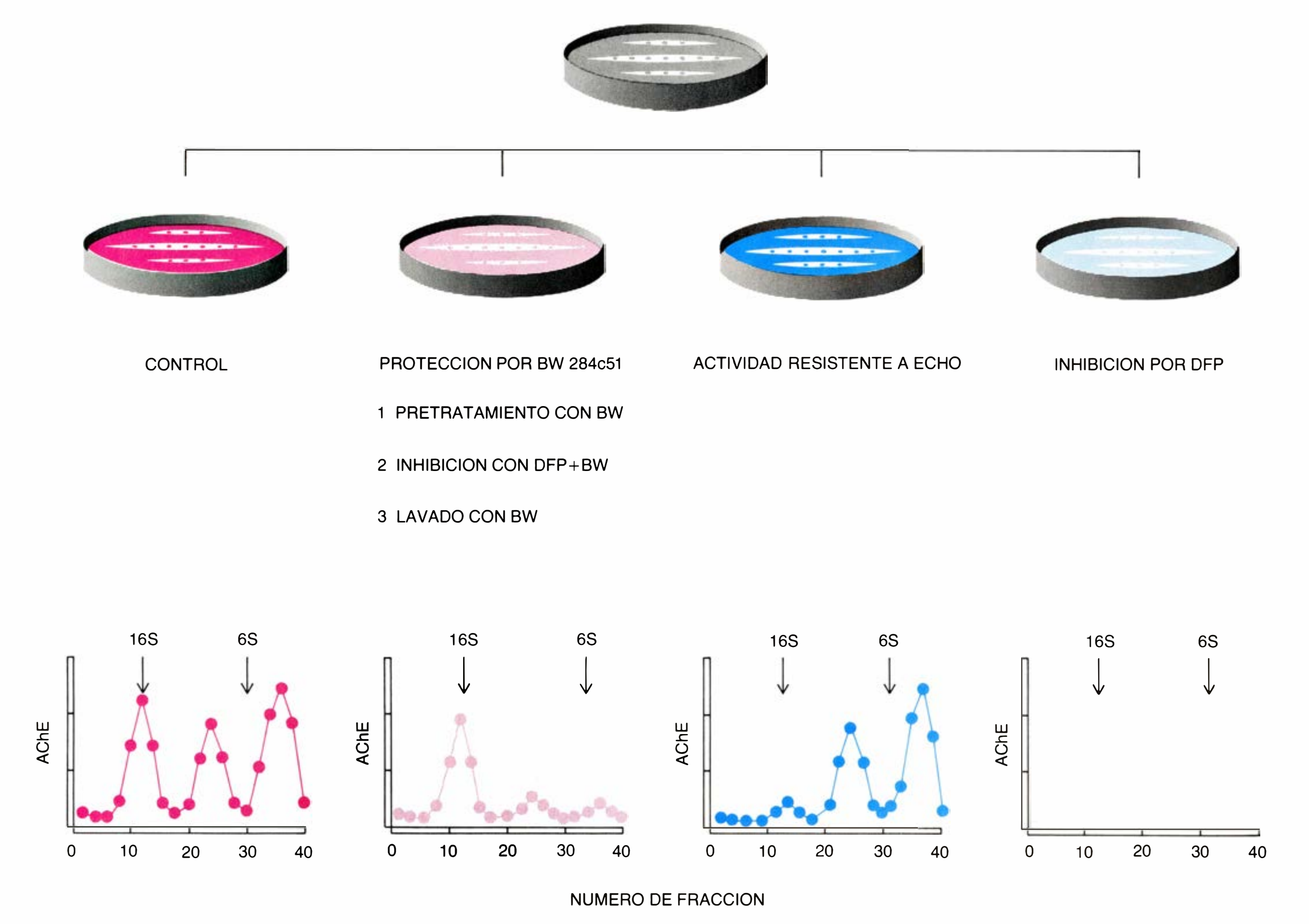
Sabiendo que la actividad colinesterásica y la forma enzimática 16*S* son características de la unión neuromuscular en un animal adulto, podemos preguntarnos si existe en los músculos en cultivo correspondencia entre la localización de la actividad colinesterásica en la superficie del miotubo y la forma 16*S*. Al igual que lo observado con los AChR, en la superficie de los miotubos existen verdaderos parches de actividad colinesterásica. Los datos muestran que estos parches corresponden a la forma sináptica 16*S*: ambos aparecen después de la fusión de los mioblastos en miotubos, ambos solubilizados de la superficie de la célula muscular por colagenasa y quedan protegidos de la inactivación con DFP (un inhibidor liposoluble e irreversible) por un inhibidor no-liposoluble reversible de la enzima, el BW284c51. Estas zonas con alta concentración de AChE semejan verdaderas uniones neuromusculares

en desarrollo, y es posible que *in vivo* sean sitios que en etapas más avanzadas del desarrollo embrionario reciban una inervación preferencial.

En el músculo adulto normalmente inervado, la AChE está concentrada en la unión neuromuscular. Sin embargo, días después de seccionar el nervio motor, la AChE disminuye más de un 50 por ciento, siendo el cambio cualitativo más importante en la forma 16*S*, la cual disminuye un 90 por ciento después de una semana de denervación. Estos resultados sugieren que el nervio regula la AChE sináptica. La forma 16*S* de la AChE reaparece cuando se reinerva el músculo esquelético, ya sea en su sitio original, la antigua unión neuromuscular, o bien en sitios alejados de la sinapsis original. La regulación de la forma sináptica de la AChE es una propiedad del nervio motor colinérgico; así, músculos lisos inervados por fibras postganglionares del simpático no contienen la forma $A_{12}$. Sin embargo, cuando los reinervan fibras motoras, las formas asimétricas $A_{12}$ son inducidas en músculos lisos, tal como lo demostramos con Bernardita Méndez y Joaquín Luco en el laboratorio de neurofisiología de la Universidad Católica de Santigo de Chile en 1978.

La importancia de la actividad muscular en la aparición de la AChE sináptica la establecieron Lomo y sus colaboradores de la Universidad de Oslo. Ellos indujeron uniones neuromusculares ectópicas implantando el nervio fibular en el músculo soleo, previamente denervado. En estas condiciones, la AChE aparece localizada en las sinapsis de neoformación. Si se cortan las fibras del nervio fibular después de 2 a 7 días de la sección del nervio soleo, no aparece la AChE. Sin embargo, si durante este período el músculo se estimula crónicamente, aparecen placas de actividad colinesterásica en zonas que antes ocupaba el nervio fibular. Estos experimentos muestran que la AChE es sintetizada por el músculo y que su aparición requiere de la actividad muscular. El hecho de que la AChE se presente en sitios donde antes había actuado el nervio indica que éste deja alguna señal para que el músculo deposite allí la AChE. Se ha sugerido que el nervio estaría induciendo una "marca" local y persistente en la lámina basal sináptica.

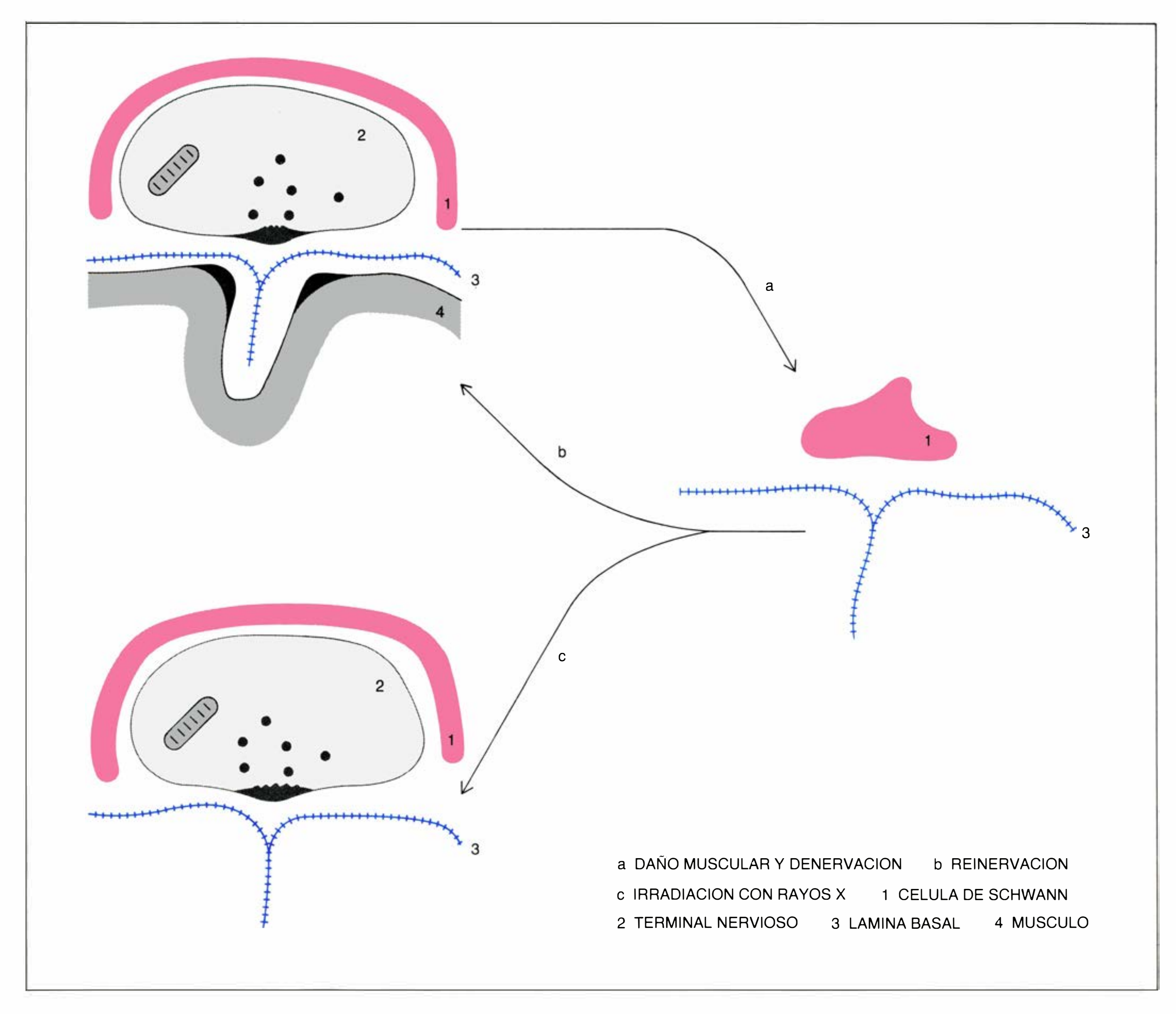
Una serie de elegantes experimentos realizados por Jack McMahan y colaboradores, primero en la Universidad de Harvard y luego en Stanford, han sugerido que la lámina basal de la unión neuromuscular difiere de la lámina basal que cubre el resto de la superficie muscular. El diseño de estos experimentos se basa en lesionar y denervar músculos de rana *in vivo*, preservando la región de inervación. Luego de este procedimiento y después de que el nervio y el músculo han degenerado, es posible reconocer los sitios sinápticos originales en las vainas de la lámina basal. Se puede evitar la reinervación eliminando grandes seg-




**6. LOCALIZACION DE LA ACETILCOLINESTERASA 16*S*. Se estableció en cultivos de células musculares diferenciadas, denominadas miotubos, mediante el uso de inhibidores de la enzima que se diferencian tanto en la calidad de su inhibición como en su permeabilidad a la membrana muscular. La existencia de la forma 16*S* extracelular se estableció usando un inhibidor reversible e impermeable a la membrana, como lo es el BW284c51, junto a un inhibidor irreversible y permeable el diisopropilfluorofosfato (DFP). La forma intracelular de la enzima AChE 16*S* se estableció usando un inhibidor irreversible e impermeable a la membrana muscular, el ecotiofato (ECHO). Las formas moleculares de la AChE se separaron por gradientes de densidad en sacarosa.**

mentos del nervio y bloqueando la regeneración del músculo por irradiación con rayos X. Así se ha examinado la lámina basal sináptica en la ausencia del nervio o del músculo. Los resultados obtenidos indican que una estructura que aparecía tan inerte es responsable de múltiples funciones, hasta entonces asociadas, en gran parte, con las células pre o postsinápticas.

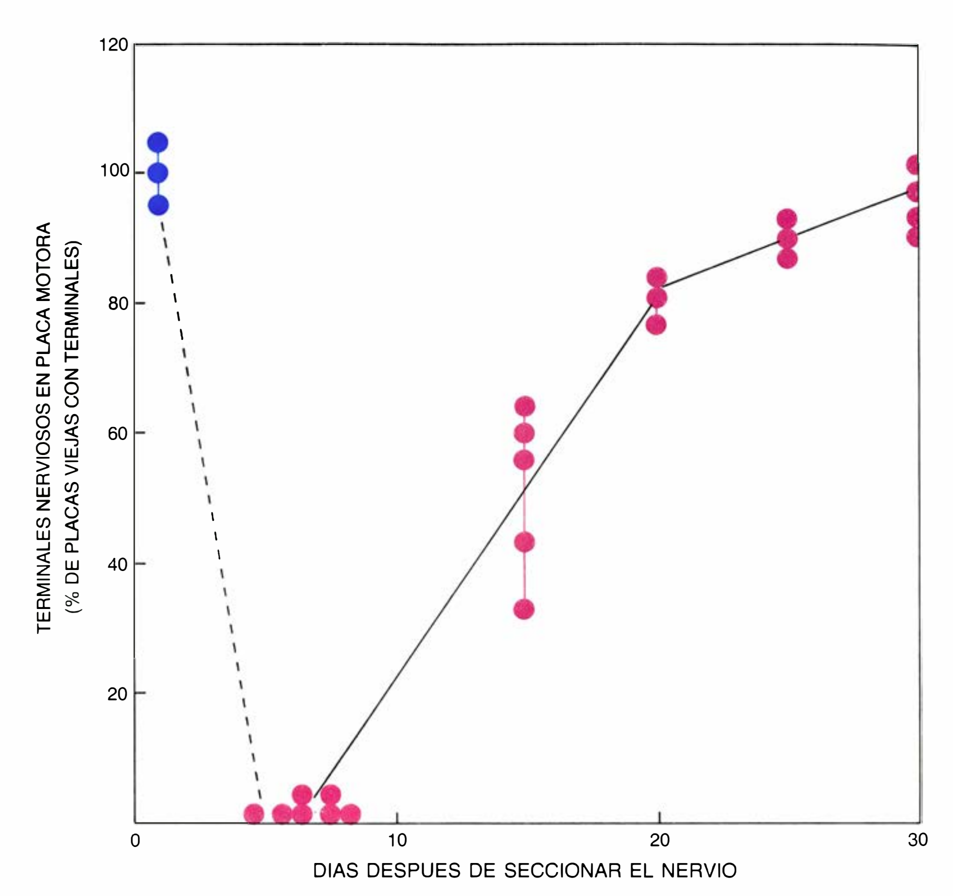
Al menos cinco diferentes funciones residen en la lámina basal sináptica: sitio de anclaje para la AChE, adhesión del nervio al músculo, especificidad topográfica durante la reinervación de un músculo denervado, diferenciación presináptica y diferenciación postsináptica. Empecemos por el sitio de anclaje para la AChE. Varios estudios sugieren que la AChE podría estar asociada con la lámina basal. El tratamiento con colagenasa, enzima que digiere material colagenoso presente en la lámina basal, libera AChE activa de la unión neuromuscular sin alterar la membrana postsináptica. La forma 16*S* de la AChE, purificada de órganos eléctricos, presenta un tallo con propiedades similares al colágeno. Como la lámina basal contiene colágeno en su forma no fibrilar, se ha sugerido que la AChE podría estar anclada en ésta por medio de su tallo. McMahan y colaboradores han realizado experimentos usando músculos en los que se han extirpado los terminales nerviosos y la fibra muscular. En estas condiciones, observaron que al menos parte de la AChE detectada histoquímicamente persiste en sitios antes inervados de la lámina basal, los cuales sobreviven a la degeneración muscular. Sin embargo, como la forma sináptica 16*S* desaparece después de la denervación, estos experimentos sólo prueban de manera indirecta la presencia de la forma 16*S* en la lámina basal de la unión neuromuscular. Experimentos realizados por el autor en la División de Neurobiología de la Universidad de California en San Francisco han mostrado que la forma 16*S*: (*1*) se encuentra localizada exclusivamente en la superficie de la célula muscular, (*2*) está asociada por interacciones iónicas al miotubo y (*3*) desaparece de la superficie del miotubo después de tratar las células con colagenasa purificada. Por tanto, podemos concluir que la lámina basal contiene a la forma sináptica de la AChE. Estu-




**7. FUNCION DE LA LAMINA BASAL SINAPTICA en la diferenciación pre y postsináptica. Quedó establecida por experimentos en los cuales se dañaba el músculo o se seccionaba el nervio y se estudiaba la regeneración de uno u otro elemento en ausencia del complementario. Así, en el ejemplo ilustrado, después de dañar el músculo y denervar el nervio, sólo quedan en la unión neuromuscular las vainas y pliegues de lámina basal y en el lugar del terminal nervioso, la célula de Schwann. Si se impide entonces la regeneración del músculo por irradiación, y se permite la regeneración, los terminales vuelven a sitios antes inervados.**

dios recientes de Enrique Brandan y el autor, en el Grupo de Neurobiología Molecular de la Universidad Católica de Chile, indican que la forma 16*S* se asocia con la lámina basal a través de azúcares complejos denominados glicosaminoglicanes, que se organizan en proteoglicanes, esto es, macromoléculas formadas por una proteína central y cadenas de glicosaminoglicanes laterales. La AChE se asocia en forma específica con proteoglicanes del tipo heparón sulfato.

De acuerdo con algunos datos, la adhesión de los terminales nerviosos a las fibras musculares podría venir mediada por componentes de la lámina basal. En efecto, al tratar músculos esqueléticos con proteasas que digieren la lámina basal, los terminales nerviosos pierden sus sitios de unión y se separan fácilmente. Esto pone de manifiesto que podría haber componentes de la lámina basal comprometidos en el mantenimiento de la integridad de la unión neuromuscular. Otras pruebas en este sentido nos las ofrecen los estudios de microscopía electrónica realizados en músculos dañados pero no denervados por Joshua Sanes y colaboradores, de la Universidad de Harvard. En estos músculos dañados los terminales nerviosos permanecen en los sitios sinápticos originales de la lámina basal varios días después de que se hayan eliminado las células musculares.




**8. LOS TERMINALES NERVIOSOS ENCUENTRAN LAS PLACAS VIEJAS denervadas en presencia de la lámina basal sináptica y en ausencia de la fibra muscular. En esta figura se muestra el proceso de regeneración desde la denervación inicial, momento en que todos los terminales nerviosos degeneran, hasta que, días después, casi todas las placas viejas denervadas se recubren de terminales nerviosos regenerados.**

Vayamos ahora con la tercera función, la relativa a la especificidad topográfica durante la reinervación de un músculo denervado. Los axones que reinervan músculos esqueléticos denervados muestran en general una clara preferencia por los sitios sinápticos originales. Aun cuando este fenómeno se conoce desde los trabajos de Ramón y Cajal, los factores que dan cuenta de tan precisa reinervación no se descubrieron hasta hace unos ocho años. Larry Marshall y colaboradores, de la Universidad de Harvard, han mostrado que no se requiere la presencia de la fibra muscular para una reinervación precisa. En efecto, los axones regeneran y entran en contacto con las vainas de la lámina basal después de que las células musculares han degenerado. En estas condiciones, cerca de un 95 por ciento de los procesos axonales que contactan las vainas lo hacen en los sitios sinápticos originales. Por tanto, se produce una reinervación topográficamente exacta, haya o no fibra muscular. Este resultado plantea la interesante posibilidad de que la lámina basal intervenga de forma destacada en el crecimiento de los axones en regeneración. Los axones podrían adherirse preferencialmente a moléculas localizadas en la lámina basal sináptica, o bien evitar moléculas ubicadas en la lámina basal extrasináptica.

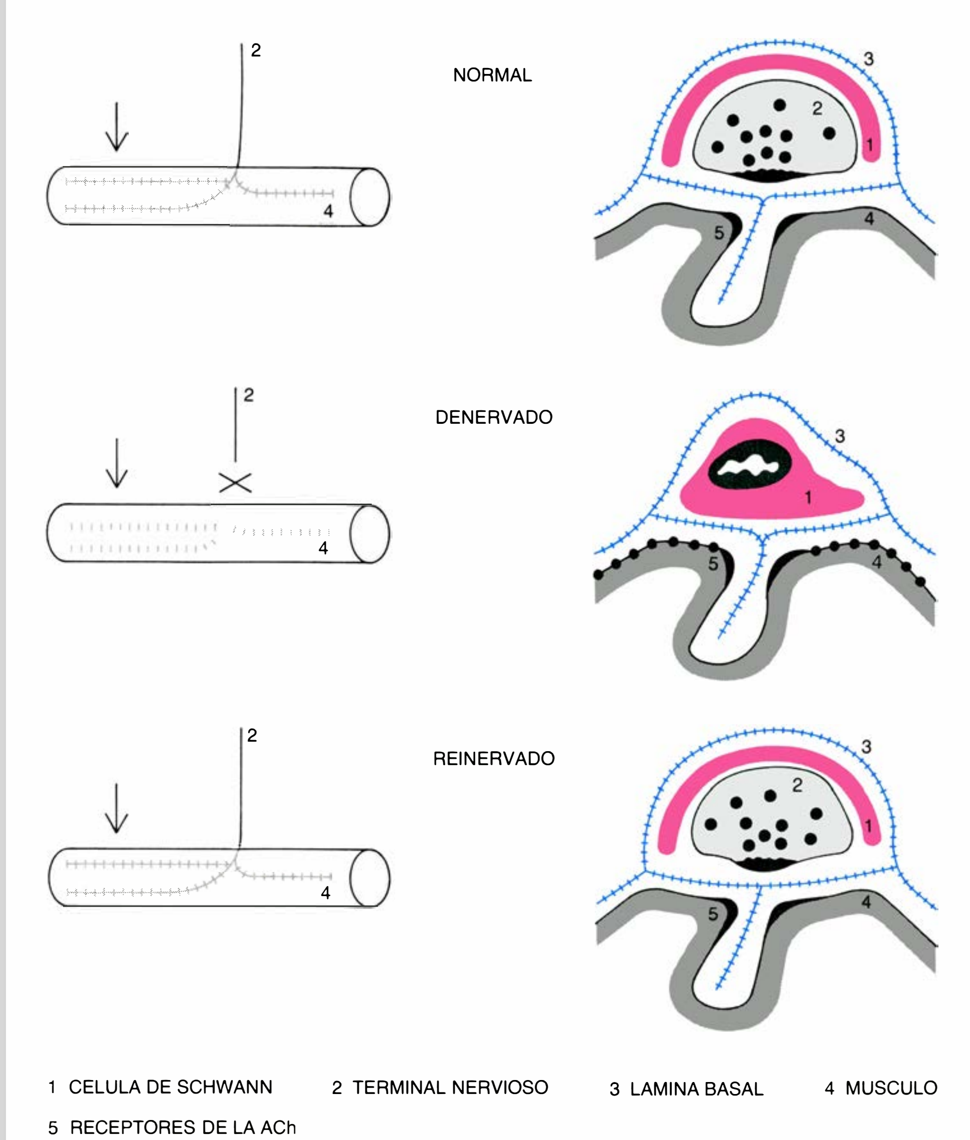
La precisión de la reinervación de un músculo esquelético se manifiesta no sólo en el crecimiento y desarrollo de los axones en regeneración, sino también en la diferenciación de éstos. De hecho, sólo las regiones axonales que se encuentran a 0,1 um de la superficie de la fibra muscular adquieren una importante acumulación de vesículas sinápticas y forman nuevas zonas activas directamente sobre las aberturas de los pliegues sinápticos. Por tanto, factores de la superficie de la fibra muscular guían la transformación de un axón en regeneración en un terminal nervioso. Sanes y colaboradores realizaron experimentos que mostraron que al menos algunos de estos factores se asocian con la lámina basal sináptica. En efecto, sólo los axones que entran en contacto con vainas de lámina basal de las fibras musculares previamente destruidas llegan a zonas activas y adquieren altas concentraciones de vesículas sinápticas. En los terminales nerviosos, las vesículas y las zonas activas se asocian preferentemente con puntos donde la lámina basal que viene de los pliegues sinápticos se encuentra con la lámina basal del espacio sináptico. De esta correspondencia específica Sanes y colaboradores han concluido que las moléculas ubicadas en la lámina basal sináptica organizan la diferenciación de los axones en regeneración en verdaderos terminales nerviosos.

Abordemos la quinta y última función de la lámina basal: la diferenciación postsináptica. Los receptores de acetilcolina se concentran en la unión neuromuscular, donde son metabólicamente estables y no presentan difusión lateral en el plano de la membrana. Es razonable suponer, por tanto, que algún elemento del exo o citoesqueleto mantiene estos parches de receptores. Steve Burden, Peter Sargent y Jack McMahan, de la Universidad de Stanford, han demostrado que ciertos factores de la lámina basal sináptica pueden inducir y mantener parches de AChE. Ellos provocaron daños y denervaron músculos esqueléticos; dejaron luego que las fibras musculares regeneraran en la ausencia de nervio. En estas condiciones, el músculo sin-

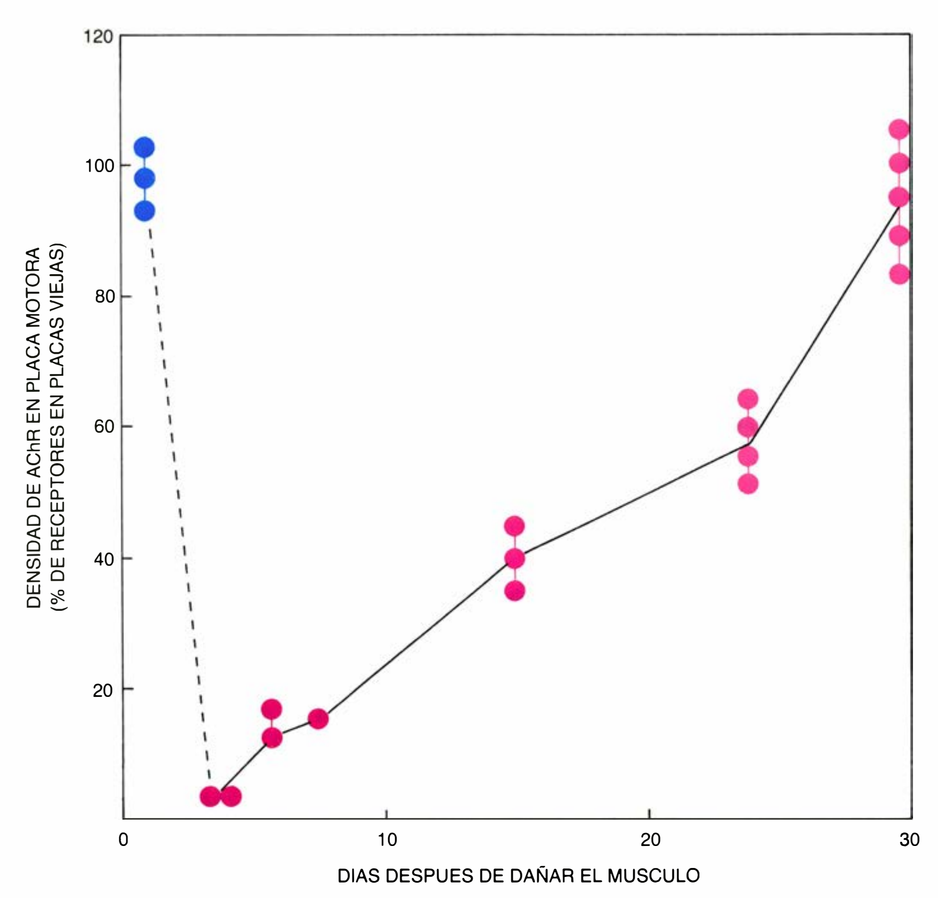
tetiza nuevos receptores, los cuales se localizan nuevamente en la región correspondiente a la lámina basal sináptica, incluso en ausencia del el nervio. De estos ensayos experimentales se concluye que los componentes de la lámina basal sináptica pueden organizar la diferenciación de la membrana postsináptica.

En general, las láminas basales son muy insolubles. Esto determina que los componentes de la lámina basal no puedan ser estudiados en su estado nativo usando las técnicas corrientes en química de proteínas. Más aún, en el caso específico de la lámina basal sináptica, ésta sólo corresponde al 0,1 por ciento del total de la lámina basal que recubre la fibra muscular, por lo que la identificación y purificación de moléculas allí presentes aparece muy difícil. Por estas razones se ha tenido que recurrir a métodos y sistemas diversos. Así, se han realizado estudios inmunocitoquímicos en músculos de animales adultos, utilizando anticuerpos generados con inmunógenos que podrían en principio estar relacionados con componentes de la lámina basal; por ejemplo, distintos tipos de colágeno. También se ha trabajado con células en cultivo capaces de sintetizar y liberar al medio moléculas sinápticas. Y, por último, con peces como *Torpedo* o *Electrophorus*, que presentan en sus órganos eléctricos grandes concentraciones de moléculas sinápticas.

Joshua Sanes, trabajando en el laboratorio de Zach Hall en la Universidad de California en San Francisco, mostró en 1979 diferencias moleculares entre la lámina basal sináptica y extrasináptica. Utilizando métodos inmu-




**9. TERMINAL NERVIOSO REGENERADO. Se diferencia normalmente en ausencia del elemento postsináptico. La reinervación no sólo es topográficamente precisa, sino que también el terminal nervioso se diferencia adecuadamente, lo que implica que las zonas activas donde se va a liberar el neurotransmisor aparecen justo frente a los pliegues de la lámina basal; aún más, en el interior del terminal las vesículas sinápticas se agrupan junto a las zonas activas frente a la invaginación de la lámina basal sináptica.**




**10. RECEPTORES DE ACETILCOLINA.** Se regeneran en ausencia del nervio, precisamente en regiones donde se encontraban antes de dañar el músculo. En estos experimentos se secciona el nervio y se impide su regeneración, permitiendo que la fibra muscular, tras el daño inicial, regenere. Lo notable es que los receptores de acetilcolina vuelven a aparecer allí mismo donde se encuentran los remanentes de lámina basal.

nocitoquímicos, estudió la unión de anticuerpos generados contra la cápsula anterior del cristalino, diferentes tipos de colágeno y AChE de cerebro. Entre los sueros producidos, cuatro de ellos reconocían antígenos sinápticos presentes en la lámina basal de la unión neuromuscular: anticápsula anterior del cristalino, antiAChE de cerebro, anticolágeno de la cápsula anterior del cristalino y anticolágeno de la membrana basal de músculos esqueléticos. Los dos últimos reconocen los antígenos sinápticos sólo después de una adsorción general con regiones de músculo libres de uniones neuromusculares. Sanes, ahora en la Universidad de Washington en Saint Louis, ha mostrado también la existencia de moléculas localizadas tanto en zonas sinápticas como extrasinápticas; tal es el caso del procolágeno tipo IV, laminina y fibronectina. Otras, por el contrario, se localizan exclusivamente en zonas carentes de uniones neuromusculares: colágeno tipo V y un nuevo tipo de colágeno presente en el tejido conectivo de músculos esqueléticos.

Junto a Laura Silberstein y Zach Hall, de la Universidad de California en San Francisco, hemos estudiado hace algunos años la aparición de antígenos sinápticos de la lámina basal en células musculares en cultivo. Uno de estos antígenos, el reconocido por el suero anticápsula anterior del cristalino, se localiza en parches en la superficie del miotubo, en zonas superpuestas a los parches de AchR; estos estudios abren la posibilidad de caracterizar bioquímicamente la estructura de un antígeno sináptico.

En conclusión, podemos decir que la lámina basal de la unión neuromuscular no sólo contiene la acetilcolinesterasa, una enzima fundamental en la fisiología muscular, sino que además es responsable de mantener la adhesión nervio-músculo. Aún más importante es la capacidad de organizar y mantener las membranas y estructuras pre y postsinápticas. Claramente, esta increíble especificidad de la lámina basal sináptica nos ofrece una nueva forma de entender y explicar problemas de la diferenciación y mantenimiento de los componentes de la unión neuromuscular. También abre la posibilidad de aprender más de esta fascinante estructura y, en el futuro, caracterizar nuevos marcadores sinápticos cuya existencia era insospechada algunos años atrás.

# La cola del campo magnético terrestre

*El viento solar arrastra al campo magnético terrestre formando una vasta cola cuya desorganización origina brillantes auroras y eyecta grandes masas de gas magnetizado al espacio interplanetario*

Edward W. Hones, Jr.

La Tierra reside en una vasta cavidad magnética. Llamada magnetosfera terrestre, la cavidad existe porque el viento solar, gas de partículas subatómicas cargadas que fluye continuamente del Sol, no puede penetrar fácilmente en el campo magnético terrestre. Antes bien, el viento solar alarga el campo formando una región más o menos cilíndrica que se extiende desde la Tierra hacia el espacio interplanetario como una manga de viento de millones de kilómetros de longitud. Esta región cilíndrica constituye la cola del campo magnético.

Nuestro conocimiento de la cola del campo magnético es el resultado indirecto de un prolongado esfuerzo por la comprensión de las luces aurorales que brillan en los cielos a altas latitudes, esfuerzo en el que se empeñaron investigadores tan notables como Galileo, Halley, Celsius y Franklin. El progreso fue relativamente lento hasta el comienzo de la era espacial, hace aproximadamente 30 años, cuando los satélites portadores de instrumentos científicos vinieron en auxilio del estudio del espacio interplanetario que rodea la Tierra. Así se llegó al descubrimiento y exploración de la magnetosfera terrestre. Las exploraciones, a su vez, condujeron a una conclusión general: el viento solar y la magnetosfera forman un gran generador eléctrico, en el que la interacción entre los campos magnéticos y el plasma (el gas de partículas del viento solar) convierte la energía cinética del movimiento del viento solar en electricidad.

Dicha energía eléctrica da lugar a una gran riqueza de fenómenos, entre los que se encuentran no sólo las bellas y enigmáticas auroras, sino también los cinturones de radiación de Van Allen, que rodean la Tierra. Otro fenómeno de éstos es el plasmoide. Descubierto en la cola del campo magnético terrestre hace sólo tres años, merced a datos aportados por satélites, consiste en enormes estructuras de plasma. Un plasmoide es un cuerpo de plasma caliente entrelazado y mantenido gracias a bucles del campo magnético. Las líneas del campo magnético que rodean a los plasmoides los impulsan desde la cola de dicho campo, cual proyectiles de cañón, a velocidades de millones de kilómetros por hora.

La interacción entre el viento solar y la magnetosfera comienza en la superficie de ésta, en la llamada magnetopausa. Su porción más cercana al Sol, donde el campo magnético terrestre detiene el movimiento hacia adelante del viento solar, se encuentra a 10 radios terrestres "corriente arriba" de la Tierra. (El radio de la Tierra, de 6370 kilómetros, utiliza habitualmente como unidad de distancia en la investigación magnetosférica.) Aquí, en la región de la Tierra que mira al Sol, el campo magnético terrestre está comprimido por la interacción con el viento solar. En el lado nocturno de la Tierra, el campo magnético se alarga lejísimos, formando la magnetocola o cola del campo magnético. El diámetro de la cola está entre 40 y 60 radios terrestres; su longitud excede el millar de ellos.

La cola del campo magnético consta de dos mitades adyacentes llamadas lóbulos, que tienen polaridades magnéticas opuestas [*véase la figura 2*]. En el lóbulo superior, o Norte, el campo magnético apunta hacia el Sol; las líneas del campo se unen a la región polar Norte de la Tierra. (La dirección del campo es la dirección hacia la cual apuntaría la punta Norte de una aguja magnética si se colocara la brújula en el campo.) En el lóbulo inferior, o Sur, el campo apunta en dirección contraria al Sol; las líneas del campo terminan en la región polar Sur. Los dos lóbulos están separados por una lámina de corriente eléctrica que fluye a través del plano central de la cola y se curva después alrededor de los lóbulos Norte y Sur. Estos bucles de corriente crean los campos magnéticos en los lóbulos.

El plasma del viento solar no queda excluido del todo de este complejo dominio magnético. En la región dirigida hacia el Sol penetra algo de plasma en la magnetosfera y puebla las regiones superficiales de la cola. Desde esas regiones, el plasma fluye a través de los lóbulos hacia el plano central de la cola, donde se concentra en la "capa de plasma". Esta es la sede de la corriente que separa los lóbulos.

Los procesos que mantienen la cola del campo magnético son las interacciones entre partículas cargadas, corrientes eléctricas, campos eléctricos y campos magnéticos. Operando conjuntamente en la cola del campo magnético, crean fenómenos físicos de suma complejidad. Tomándolas aisladamente, sin embargo, cada tipo de interacción es bastante sencilla. Así, el conocimiento de las interacciones in-

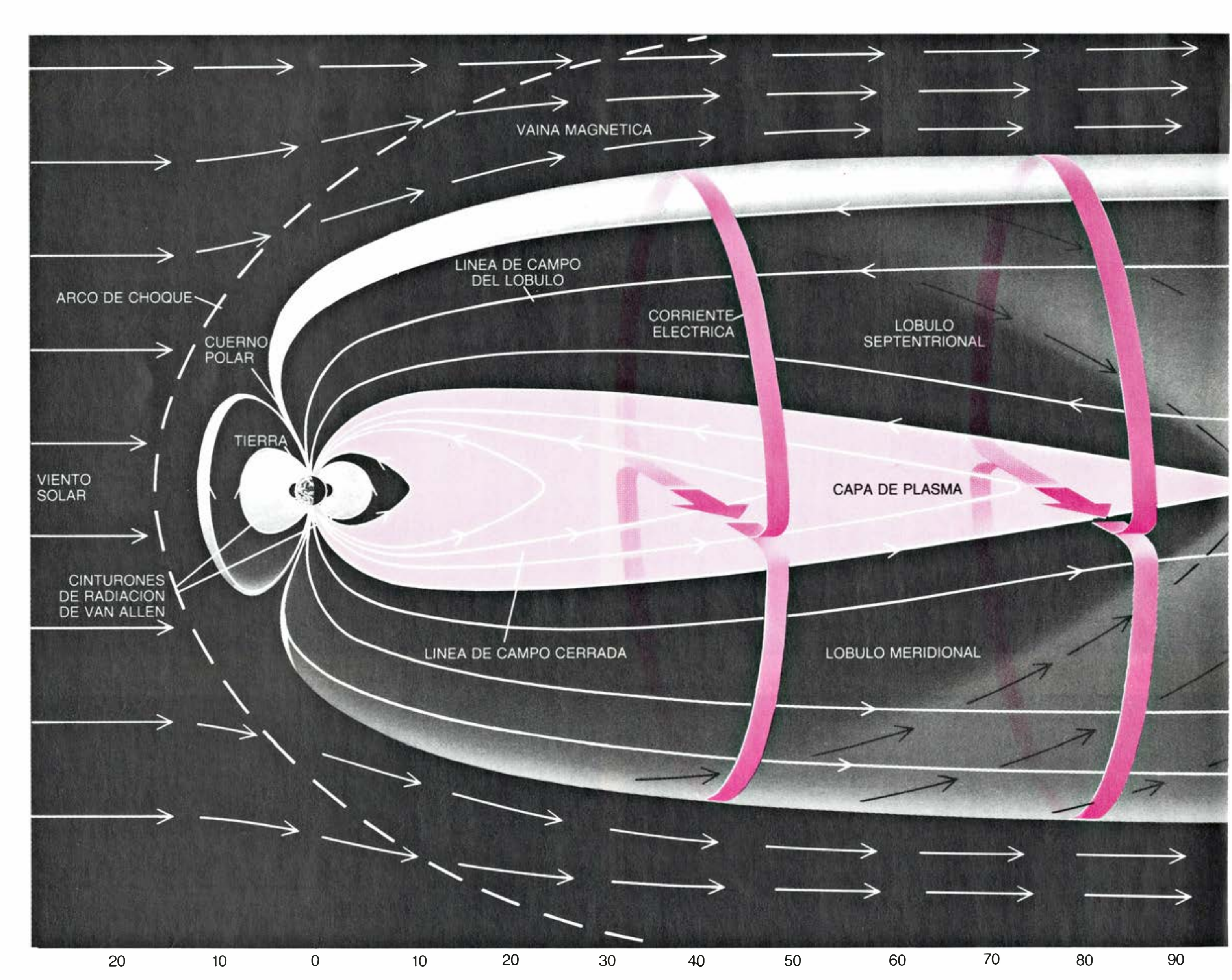
**1. ENERGIA DE LA COLA DEL CAMPO MAGNETICO liberada en la ionosfera, a unos 100 kilómetros por encima de la superficie de la Tierra. Es responsable de la aurora boreal que se muestra *(de arriba abajo y de izquierda a derecha)* en esta sucesión de fotografías que Robert H. Eather, del Boston College, realizó cerca de media noche en Churchill, en la bahía de Hudson, al norte de Manitoba, cierto día del mes de marzo. Eather instaló en su cámara un objetivo de ojo de pez, de modo que cada imagen recogiera la totalidad del cielo, de horizonte a horizonte; en todas las imágenes, el horizonte Sur está en la parte inferior. Las auroras se producen por bombardeo de los átomos en la ionosfera. Los electrones, expulsados del Sol en forma de viento de partículas cargadas (el llamado viento solar), quedan aprisionados en la cola del campo magnético y se acercan a la Tierra cuando una "subtormenta magnética" disgrega la configuración normal de la magnetosfera. Estas imágenes cubren un intervalo de 20 minutos durante una de tales subtormentas.**




dividuales puede servir para dar idea de la estructura que su acción conjunta crea.

Un campo magnético actúa sobre una partícula cargada mediante la fuerza de Lorentz, que es proporcional a la intensidad del campo magnético y a la componente de la velocidad de la partícula perpendicular al campo; la fuerza está dirigida formando ángulo recto con ambas [*véase el esquema superior de la figura 3*]. En un campo magnético uniforme, la fuerza de Lorentz determina que los electrones y protones describan circunferencias en sentidos opuestos. Si el movimiento original de una partícula tenía una componente paralela o antiparalela a la dirección del campo, la combinación con el movimiento circular confiere a la trayectoria una forma helicoidal. Al moverse la partícula, el centro de su trayectoria helicoidal describe una línea de fuerza del campo magnético. En efecto, las partículas cargadas están "ligadas" a las líneas del campo magnético: cada partícula traza una hélice alrededor de una línea.

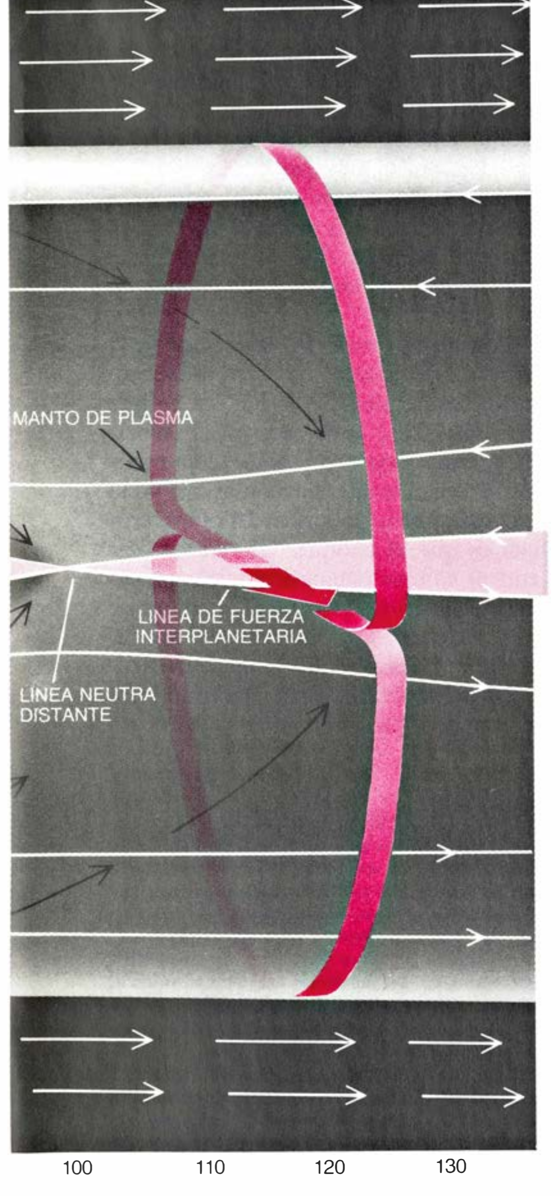
Una segunda clase de interacción entre partículas y campos es la llamada interacción E-B. Cuando existe un campo eléctrico situado perpendicularmente a un campo magnético [*véase el esquema central de la figura 3*] las partículas en movimiento ligadas a las líneas del campo magnético experimentan un corrimiento perpendicular al campo eléctrico (E) y al campo magnético (B). Las partículas positivas –así los protones– y las negativas –los electrones, por ejemplo– se desvían en la misma dirección y a la misma velocidad; de aquí que las partículas de un plasma (un gas con igual número de partículas positivas que negativas) se muevan todas conjuntamente. Se puede pensar que las líneas del campo




**2. CORTE ESQUEMATICO DE LA MAGNETOSFERA que muestra sus principales características magnéticas y eléctricas. El viento solar se desvía a lo largo de una línea llamada arco de choque. Pasa, después, en torno a la magnetosfera en una región llamada vaina magnética. La cola del campo magnético es la parte de la magnetosfera "corriente abajo" respecto a la Tierra; su superficie se llama magnetopausa. En su lóbulo superior, o Norte, las líneas de fuerza del campo magnético apuntan hacia el Sol; en el lóbulo inferior, o Sur, las líneas de fuerza apuntan en sentido contrario. Entre los lóbulos se extiende una capa de plasma (gas de partículas eléctricamente cargadas), a lo largo de la cola, transportando una corriente eléctrica transversal *(en color)* que circunda los lóbulos. A unos cien radios terrestres corriente abajo de la Tierra, se encuentra una línea neutra magnética que también corta la cola (el dibujo muestra sólo su punto medio). En la línea neutra confluyen y se reconectan las líneas de campo de los lóbulos, transportadas hacia la capa de plasma por fuerzas electromagnéticas; la reconexión convierte las líneas de los lóbulos en líneas cerradas del campo (bucles ligados a la Tierra) y en líneas de fuerza interplanetarias que se**

magnético se mueven con el plasma a una velocidad igual a la velocidad E-B de las partículas. Reunidas esas condiciones, se dice que las líneas del campo magnético están congeladas en el plasma.

Hay circunstancias bajo las cuales las líneas del campo magnético se debilitan mucho o cambian muy bruscamente en el espacio o en el tiempo; entonces, las partículas de plasma y las líneas del campo ya no se mueven juntas. Antes bien, las líneas del campo se "difunden" en el plasma. Bajo tales condiciones, las líneas de campo de sentidos opuestos pueden unirse, romperse y reconectarse en una nueva combinación. Esta "reconexión magnética", así se llama, es crucial en la magnetosfera, tanto en el proceso de adquisición de energía a costa del viento solar (proceso más o menos continuo) cuanto en el de su liberación de energía a costa de la cola del campo magnético, creando auroras y plasmoides.




**extienden hacia fuera, de modo tal que ambos extremos se funden en el viento solar. El plasma de la vaina magnética que penetra en los cuernos polares ocupa la región de los lóbulos llamada manto de plasma. El plasma deriva hacia la capa de plasma y se inyecta en ella mediante el proceso de reconexión. La cola se extiende en el espacio mil radios terrestres, por lo menos. (Los dibujos son obra de Ian Worpole.)**

La reconexión magnética tiene lugar cuando se unen regiones de campos magnéticos opuestos. Puede visualizarse fácilmente por medio de las variaciones de los campos magnéticos que se producen cuando dos imanes de herradura de idéntico tamaño e intensidad se ponen juntos de tal manera que se alineen los polos opuestos, es decir, que el polo norte de una de las herraduras se enfrente al polo sur de la otra. Los campos magnéticos producidos por los imanes varían, de modo que algunas de las líneas de fuerza que originariamente se dirigían del polo norte al polo sur de uno de los imanes, van ahora del polo norte de uno al polo sur del otro.

La reconexión magnética en la magnetosfera terrestre es más complicada y menos fácil de explicar intuitivamente. La complejidad se produce porque las regiones magnéticas de la magnetosfera son plasmas, fluidos que conducen fácilmente la corriente eléctrica. Las corrientes crean campos magnéticos que, a su vez, dan nuevas formas a los plasmas. No obstante, el proceso básico es el mismo: la reconexión magnética tiene lugar en la magnetosfera terrestre cuando se unen dos regiones de plasma de magnetización opuesta.

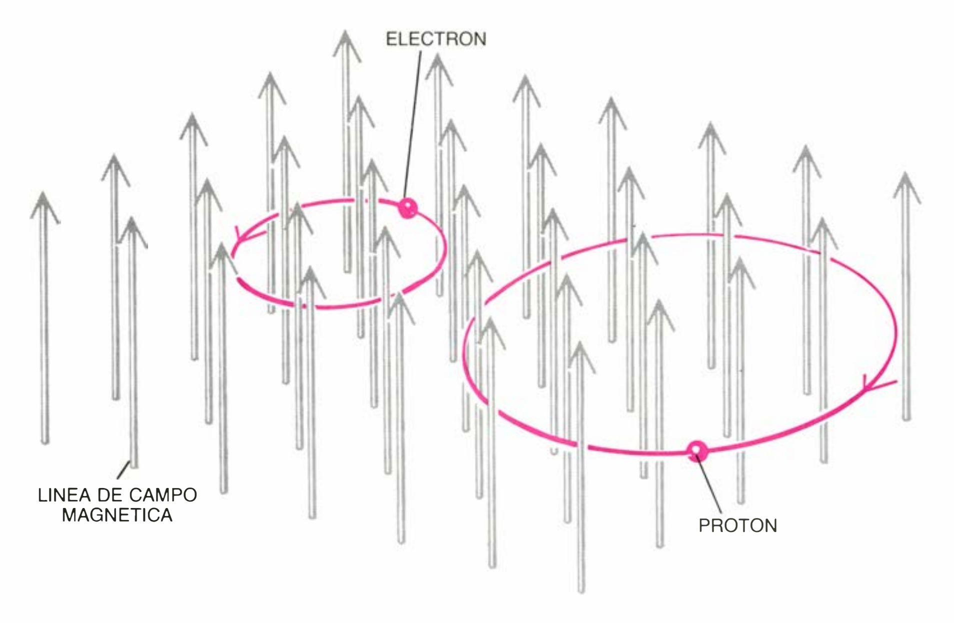
Una zona para la reconexión magnética en la magnetosfera es la lámina o capa de plasma. De hecho, la reconexión renueva continuamente la lámina de plasma. El proceso comienza cuando un plasma del viento solar logra entrar en la magnetosfera, en lugares tales como los cuernos o antenas polares (donde las líneas del campo magnético de la magnetosfera se inclinan hacia la Tierra sobre los polos magnéticos Norte y Sur). Inicialmente, el plasma fluye corriente abajo en la magnetosfera, a lo largo de las líneas del campo magnético en los niveles exteriores de los lóbulos Norte y Sur, región llamada manto de plasma. Al fluir, el plasma queda sometido al corrimiento E-B, que lo arrastra hacia la lámina de plasma, en el plano central de la cola. Los plasmas que fluyen transportan consigo campos magnéticos opuestos y congelados. En algún punto a lo largo de la interfase entre los lóbulos, la condición de campo congelado se rompe, produciéndose una región de difusión, es decir, una región en la que los campos magnéticos se pueden mover a través del plasma.

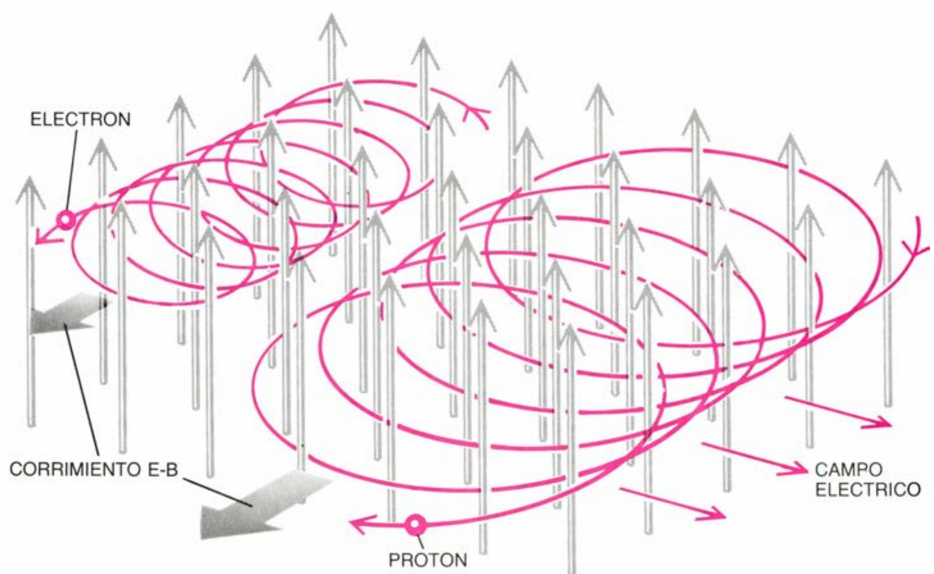
Dentro de la región de difusión, las líneas de campo de sentidos opuestos se unirán y se reconectarán a través de la capa de plasma. (Las líneas que apuntan hacia el Sol provienen del lóbulo Norte; las dirigidas al punto antisolar, del lóbulo Sur.) En el lugar donde se tocan las líneas de campos opuestos, la intensidad magnética neta del campo es nula, debido a la presencia de magnetizaciones opuestas. De aquí que el lugar de reconexión se llame línea magnética neutra. Normalmente está situada en la capa de plasma, a unos 100 radios terrestres corriente abajo de la Tierra. A cada lado de la línea neutra, el plasma se está expulsando de la región de difusión a lo largo del plano central de la cola del campo magnético, en forma de un delgado chorro en cuña. El chorro expulsado hacia la Tierra renueva la capa de plasma; el chorro expulsado hacia la cola fluye corriente abajo, nuevamente hacia el viento solar.

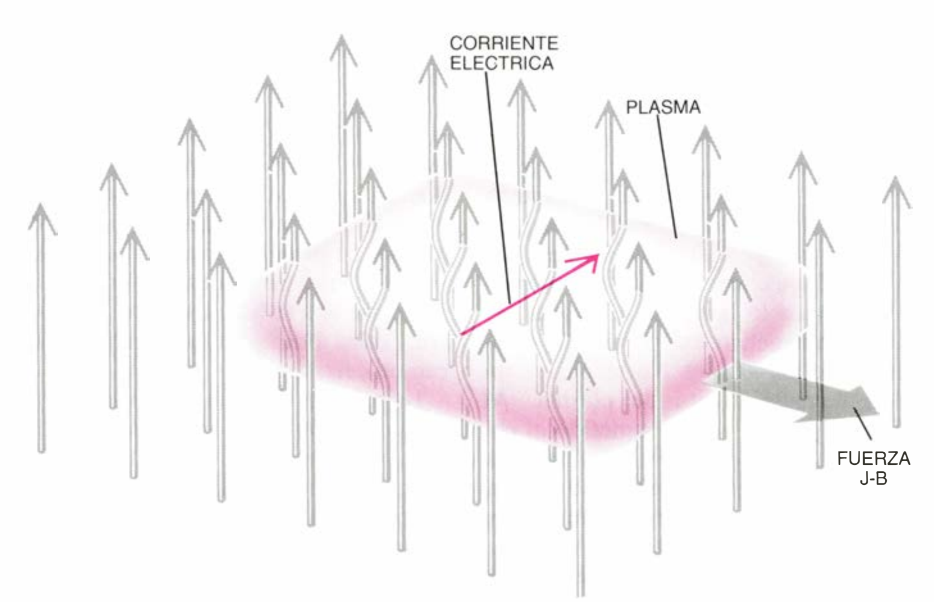
Como resultado de la reconexión, la cola del campo magnético tiene tres tipos de líneas de campo. (Un cuarto tipo, que abordaremos luego, sólo existe de modo transitorio y forma la estructura magnética de los plasmoides.) Cada tipo, que ocupa una porción particular de la cola, está caracterizado por su relación con la Tierra. El primero de ellos lo integran las líneas de campo de los lóbulos. Un extremo de cada una de esas líneas está unido a la Tierra; el otro se extiende corriente abajo hacia el viento solar. Tales líneas se llaman líneas de campo abiertas.

El segundo tipo se encuentra en la capa de plasma del lado próximo a la Tierra de la línea magnética neutra. Aquí, cada línea de campo incluye los extremos situados hacia la Tierra de dos líneas lobulares que se han reconectado. Así, cada línea de campo de la capa de plasma hacia el lado terrestre de la línea neutra es un bucle cuyos dos extremos se hallan unidos a la Tierra. Tales líneas se llaman líneas de campo cerradas.

La tercera clase se encuentra en la capa de plasma corriente abajo de la línea magnética neutra. Aquí, cada línea de campo comprende los extremos corriente abajo de dos de las líneas de campo lobulares que se han reconectado; son así bucles completamente libres de la Tierra. Sus extremos se ale-




jan, en dirección contraria al Sol, hacia el espacio interplanetario. Se conocen como líneas de fuerza interplanetarias.

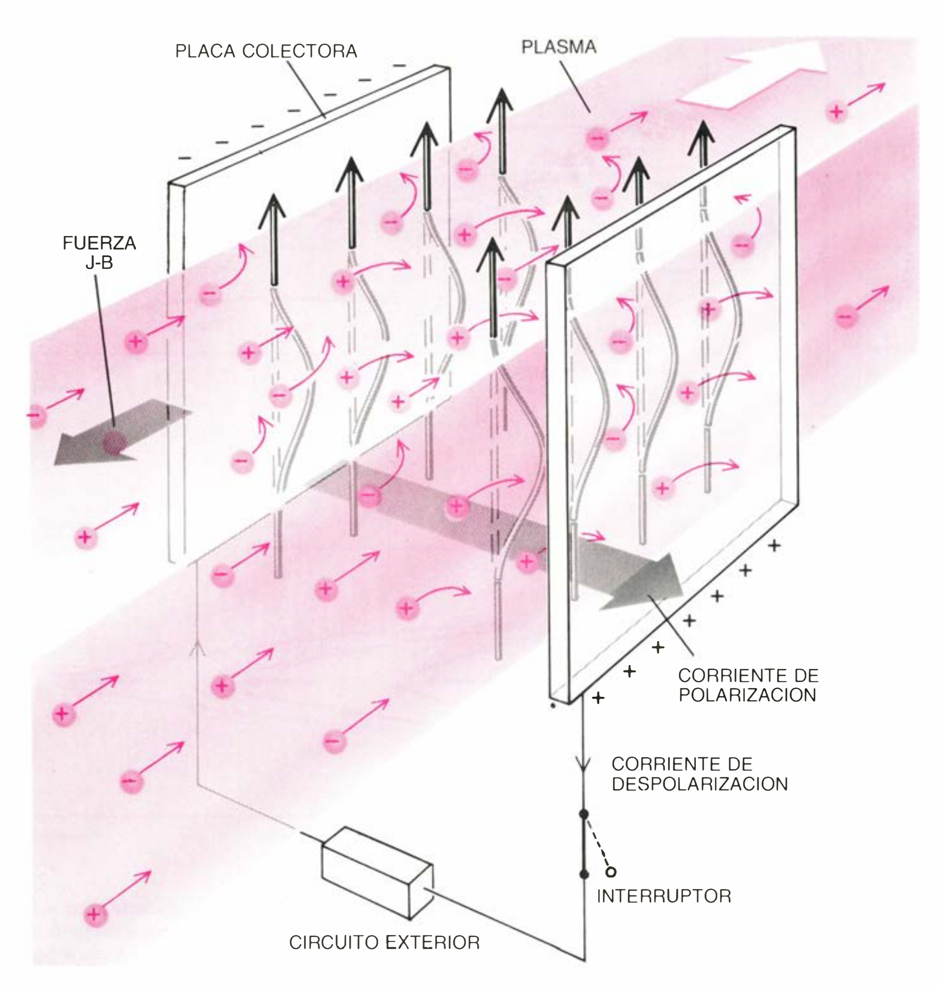
Esto no quiere decir que la cola del campo magnético mantenga siempre una configuración magnética limitada a esos tres tipos de líneas de campo magnético. Hay veces en que la interacción entre el viento solar y el campo magnético terrestre sobrecarga de energía a la cola del campo magnético, dando lugar al fenómeno llamado subtormenta magnética, que distorsiona la configuración del campo. La subtormenta es el mecanismo mediante el cual la magnetosfera libera intermitentemente grandes cantidades de energía que ha sido acumulada en la cola del campo magnético. Parte de esta energía se emplea en crear las auroras, cerca de la Tierra, mientras que el resto se lanza, corriente abajo, al viento solar en forma de un plasmoide. (Una subtormenta, que dura alrededor de una hora, difiere de una tormenta geomagnética, que dura al menos un día y se produce cuando una fulguración solar inicia una onda de choque en el viento solar.)

Para comprender la subtormenta, la aurora y los plasmoides, es preciso conocer los detalles de la manera en que el viento solar cede energía al campo magnético terrestre y cómo esta energía alarga el campo hasta más allá de lo que éste puede soportar. El punto crucial es que el choque del viento solar con el campo magnético terrestre genera electricidad porque el viento solar y la magnetosfera constituyen un generador magnetohidrodinámico (MHD).

En un generador MHD construido por el hombre, un electroimán genera un

**3. TRES FUERZAS contribuyen a crear y mantener la estructura de cola de la magnetosfera. En un campo magnético uniforme, la fuerza de Lorentz *(arriba)* hace que los electrones y protones describan circunferencias en sentidos opuestos, con lo que las partículas quedan “ligadas” al campo magnético. Si se superpone al campo magnético otro eléctrico y perpendicular, las partículas cargadas adquieren, además, un movimiento llamado corrimiento E-B *(centro)*. Este corrimiento arrastra los centros de las trayectorias circulares en una dirección perpendicular a la de ambos campos. De esta manera, el plasma de los lóbulos de la cola del campo magnético deriva hacia la capa de plasma. Finalmente, el plasma que transporta una corriente eléctrica que fluye perpendicularmente a un campo magnético dado experimenta una fuerza J-B *(abajo)*. Esta fuerza acelera el plasma en una dirección perpendicular tanto a la de la corriente como a la del campo magnético (regla de la mano derecha). La propia corriente origina un campo magnético que deforma el campo magnético original, curvando sus líneas en dirección opuesta a la de la fuerza. La fuerza J-B, que se opone al flujo del plasma del viento solar, deforma el campo magnético terrestre y crea la cola.**

intenso campo magnético entre dos placas, mientras que el plasma fluye entre ellas en dirección perpendicular al campo magnético [*véase la figura 4*]. El dispositivo que recoge la energía producida por el generador es un circuito externo que conecta las placas. Cuando las partículas penetran en el campo, la fuerza de Lorentz las desvía; los electrones se mueven hacia una de las placas y los protones hacia la otra, creando dentro del plasma una migración de cargas eléctricas llamada corriente de polarización. Si las partículas tuvieran suficiente espacio, la desviación las haría describir trayectorias circulares. Aquí intervienen, sin embargo, las placas conductoras. Una placa recoge la carga negativa, mientras que la otra recoge la positiva. La carga abandona el generador y se mueve a lo largo del circuito externo constituyendo la llamada corriente despolarizadora.



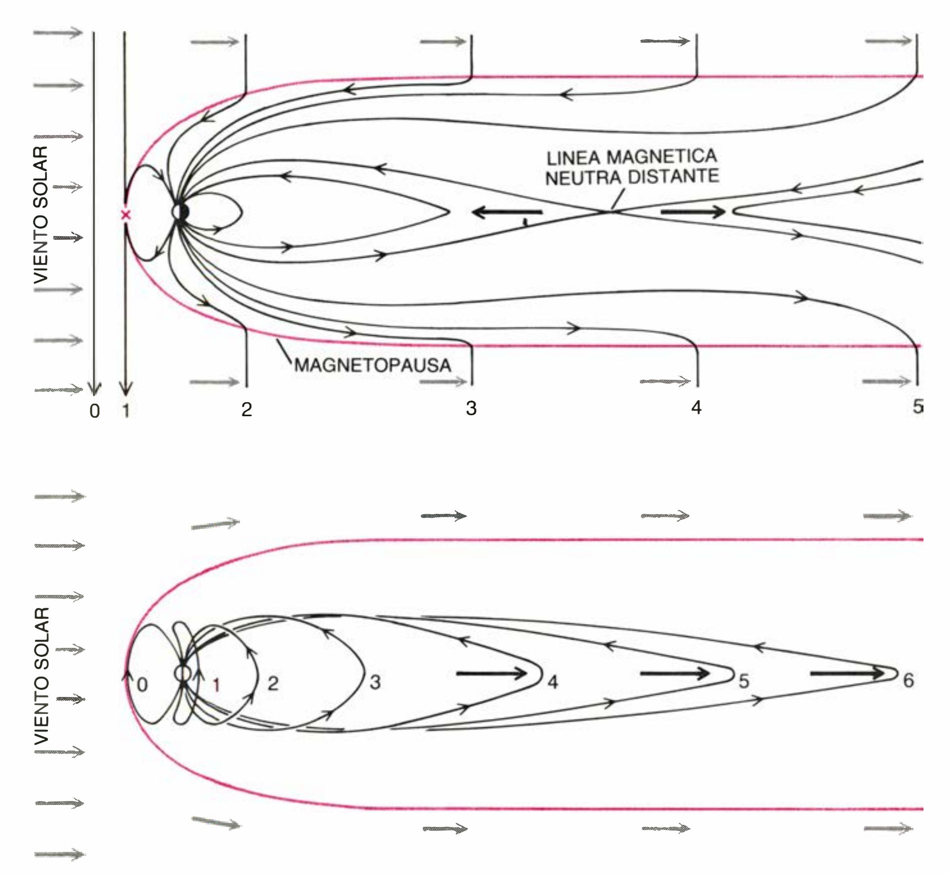

**4. GENERACION MAGNETOHIDRODINAMICA de energía eléctrica a partir de la energía cinética de un plasma que se mueve: tal es el mecanismo mediante el cual el viento solar transfiere energía a la cola del campo magnético. En un generador magnetohidrodinámico, el plasma *(en color)* se encuentra con un campo magnético; la fuerza de Lorentz impulsa a las partículas del plasma en sentidos opuestos (dependiendo de su carga) hacia las placas que están a ambos lados del campo. La emigración constituye una corriente eléctrica *(corriente de polarización)* perpendicular al campo magnético; cuando se cierra el interruptor del circuito externo *(abajo)*, circula corriente eléctrica. La fuerza J − B (*flecha de trazo grueso*) se opone al movimiento del plasma; el campo magnético causado por la corriente de polarización deforma el campo original.**

En la magnetosfera, el proceso magnetohidrodinámico genera electricidad en dos regiones diferentes. El generador más importante surge como consecuencia de la reconexión magnética en el límite de la magnetopausa situado en dirección al Sol. Allí el viento solar se encuentra de frente con el campo magnético terrestre. El propio viento incorpora un campo magnético, formado por líneas de fuerza que se alejan del Sol. La dirección hacia la que apuntan las líneas de fuerza cambia con el tiempo, al parecer de forma aleatoria. Cuando las líneas de campo apuntan hacia el Sur, esto es, en dirección opuesta a la de las líneas del campo terrestre en la parte de la magnetopausa dirigida hacia el Sol, se pueden reconectar con las del campo magnético terrestre y quedar así ligadas a la Tierra [*véase la figura 5*]. El plasma del viento solar, fluyendo a lo largo de este campo ligado a la Tierra, completa entonces los elementos básicos de un generador MHD. La ionosfera terrestre (capa ionizada y conductora de la electricidad de la atmósfera, situada a unos 100 kilómetros de altura sobre la Tierra) constituye el circuito externo de baja resistencia a lo largo del cual puede moverse la carga de polarización del viento solar.

Cuando el campo del viento solar apunta hacia el Norte, cosa que hace aproximadamente durante la mitad de tiempo, baja la reconexión magnética en la parte de magnetosfera dirigida hacia el Sol, queda muy reducida y disminuye el flujo de energía eléctrica. Incluso entonces, el segundo generador, menos potente, persiste en su actividad. A bajas latitudes, parte del plasma del viento solar fluye a través de las líneas cerradas del campo magnético terrestre. No se sabe cómo penetra, pero no parece depender demasiado de la dirección del campo magnético del viento solar. En cualquier caso, la entrada crea de nuevo las condiciones para la generación magnetohidrodinámica de energía: un plasma que se mueve en ángulo recto respecto a un campo magnético, con un camino de descarga de la polarización a través de la ionosfera.

Mediante esas posibilidades de generación magnetohidrodinámica, el viento solar inyecta electricidad en el campo magnético terrestre. A su vez, la corriente eléctrica alarga las líneas del campo terrestre. El alargamiento resulta de un fenómeno electromagnético conocido con el nombre de regla de la mano derecha [*véase el esquema inferior de la figura 3*]. La corriente eléctrica (en este caso la corriente producida por el proceso magnetohidrodinámico) que fluye en el plasma perpendicularmente al campo magnético, ejerce sobre el plasma una fuerza perpendicular a la corriente (J) y al campo magnético (B). Al mismo tiempo, deforma las líneas del campo magnético en dirección opuesta a la de la fuerza.

La mayor o menor deformación de las líneas de fuerza en el generador MHD depende de la intensidad del campo magnético, la intensidad de la corriente de polarización y el propio tamaño del generador. En los generadores MHD que se están desarrollando en la Tierra para generar corriente, la deformación es despreciable. En la magnetosfera, en cambio, es prodigiosa. Si el viento solar se detuviera, dejando de tener efecto el intercambio de energía, el campo magnético terrestre adoptaría la configuración simétrica de un campo dipolar: el campo de una




**5. DOS SEDES de generación magnetohidrodinámica surgen de la interacción entre el viento solar y el campo magnético terrestre. Cada dibujo muestra posiciones sucesivas de una línea del campo magnético. En el dibujo superior (corte vertical medio de la magnetosfera), la línea de fuerza, que apunta hacia el Sur *(O)*, forma parte inicialmente del campo magnético transportado desde el Sol por el viento solar. En el límite de la magnetopausa que mira hacia el Sol *(color)*, la línea del campo encuentra una de las líneas del campo terrestre y se reconecta con ella *(1)*, ligándose así a la Tierra. La porción exterior de la línea deriva corriente abajo *(2-5)* con el viento solar; la porción interior proporciona un camino a la ionosfera a lo largo del cual pueden moverse las cargas de polarización generadas por procesos magnetohidrodinámicos. En el dibujo inferior (vista lateral de la magnetosfera), una línea de campo en el borde de la magnetosfera *(O)* deriva corriente abajo *(1-6)* con el plasma del viento solar que ha atravesado la superficie anterior de la magnetosfera. El plasma, al atravesar líneas de campo como éstas, crea de nuevo los elementos de un generador magetohidrodinámico en el que las cargas de polarización hallan un camino hacia la ionosfera.**

barra imantada. En vez de ello, y como consecuencia de la fuerza producida por la interacción entre plasma y campo magnético, el lado del campo opuesto al Sol se convierte en la cola del campo magnético, distribución magnética mucho más extensa que representa el almacenamiento de una gran cantidad de energía: unos $10^{11}$ megajoule, cantidad de energía eléctrica que Estados Unidos tarda en consumir varios días. El ritmo al que los dos mecanismos magnetohidrodinámicos suministran energía a la magnetosfera está entre los $10^5$ y los $10^6$ megawatt, aproximadamente la tasa de consumo de energía en los Estados Unidos.

El almacenamiento de parte de esta energía en la magnetosfera interna, cerca de la Tierra, da lugar a la producción de intensas auroras; además, inyecta iones y electrones de alta energía en órbitas situadas a distancias de entre uno y seis radios terrestres, poblando así los cinturones de radiación de Van Allen. Además, gran parte de la energía y del plasma almacenados en la cola del campo magnético vuelven finalmente al viento solar en forma de plasmoides.

Las fotografías que abarcaban la totalidad del cielo, realizadas a intervalos de un minuto en una red de observatorios árticos durante el Año Geofísico Internacional (1957-58), han contribuido en gran manera a establecer la porción terrestre de esta sucesión de acontecimientos interconectados. Según Syun-Ichi Akasofu, de la Universidad de Alaska en Fairbanks, que dedujo su modelo de las fotografías, una subtormenta auroral (es decir, el aspecto de la subtormenta magnetosférica que es visible desde las regiones polares de la Tierra) comienza con la brusca iluminación de un arco auroral. La aurora se extiende después hacia el Este, hacia el Oeste y hacia el Polo, alcanzando las latitudes magnéticas de entre 75 y 80 grados magnéticos Norte y Sur en una media hora. Seguidamente comienza la fase de recuperación de la subtormenta, que dura alrededor de una hora; mientras tanto, la aurora se debilita hacia el Ecuador. Las observaciones realizadas conjuntamente en las regiones polares Norte y Sur muestran que la subtormenta evoluciona simultáneamente en ambos hemisferios.

Los fenómenos magnetosféricos, de los cuales la aurora es una parte terrestre, se conocen ahora, al menos parcialmente. De hecho, hacia principios de los años setenta, los investigadores de la Universidad de California en Los Angeles, entre ellos Ferdinand V. Coroniti, Charles F. Kennel, Robert L. McPherron y Christopher T. Russell, habían desarrollado un modelo teórico para la causa de las subtormentas magnetosféricas que sigue siendo la base de las ideas hoy dominantes. El modelo, llamado de la línea neutra, desarrolla ideas presentadas por James W. Dungey, de la Universidad de Cambridge, a principios de los sesenta. El propio trabajo del autor, que confirma y extiende el modelo, se basa principalmente en medidas del plasma de la cola del campo magnético tomadas desde satélites, mediante instrumentos diseñados por Samuel J. Bame, Sidney Singer y otros colegas míos del Laboratorio Nacional de los Alamos. Un aspecto notable del modelo es que predice la formación y expulsión de plasmoides.

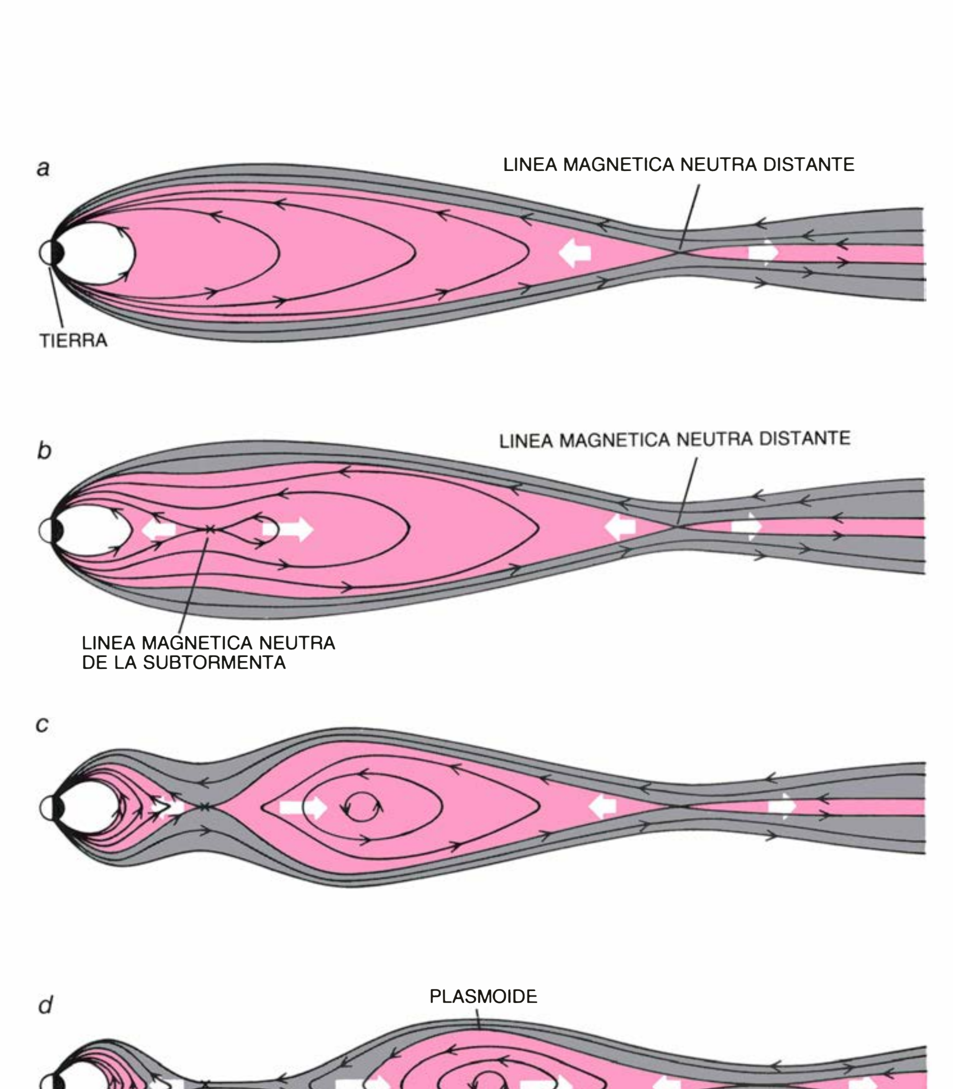
En el instante en que se inicia una subtormenta auroral en la Tierra, se forma espontáneamente en la capa de plasma, a unos 15 radios terrestres corriente abajo de la Tierra (aproximadamente una cuarta parte de la distancia de la Tierra a la órbita de la Luna), una nueva línea neutra, llamada línea neutra de la subtormenta o línea neutra cuasi-terrestre [*véase la figura 6*]. La formación de la nueva línea neutra va generalmente precedida de un extremado alargamiento de las líneas magnéticas, hasta distancias por encima de los siete radios terrestres corriente abajo de la Tierra. El alargamiento, que se puede formar en una hora aproximadamente, es consecuencia de un crecimiento del ritmo de reconexión de líneas de campo en la magnetopausa y, consecuentemente, de un aumento del ritmo de cesión de energía del viento solar a la cola del campo magnético por medio de la magnetohidrodinámica.

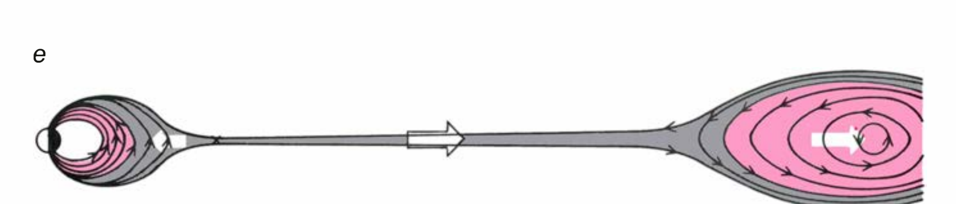
Por su parte, la formación de la

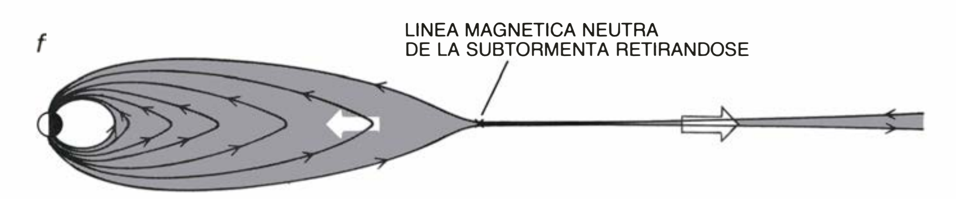
nueva línea neutra distorsiona la corriente eléctrica que cruza la capa de plasma en el lado de la línea próximo a la Tierra. Como resultado, parte de la corriente que se manifiesta a través de la cola del campo se reduce bruscamente, de suerte que las líneas del campo magnético en la región se acortan, a su vez, de una forma drástica y se contraen, o "colapsan", hacia la Tierra, volviéndose más parecidas a las del campo dipolar. El brusco colapso desencadena una lluvia de electrones en la alta atmósfera, produciendo auroras cuando los electrones bombardean los átomos de la atmósfera a unos 100 kilómetros de altura sobre la superficie de nuestro planeta.

Mientras tanto, han comenzado los sucesos que conducen a la formación de un plasmoide. Las líneas de fuerza alargadas que confluyen en la nueva línea neutra comienzan a reconectarse. Forman líneas de campo cerradas y más cortas, de la línea neutra hacia la Tierra (el colapso de líneas de fuerza mencionado arriba), y líneas de campo cerradas, de la línea neutra a la cola. Estas últimas abarcan la distancia desde la línea neutra de la subtormenta hasta la línea neutra distante y anterior a la tormenta. La proyección de plasma desde la región de la reconexión arrastra los bucles cerrados hacia la cola y las líneas de fuerza acortadas hacia la Tierra, a velocidades de centenares de kilómetros por segundo.

Cuando se han completado las reconexiones de las líneas de fuerza cerradas de la lámina de plasma, queda desligado de la Tierra un conjunto de bucles cerrados, cada uno interior a otro. La longitud de esta configuración (en el momento de formarse) es de 70 a 80 radios terrestres. Su anchura puede ser la mitad de la anchura de la cola, es decir, de 20 a 25 radios terrestres. Su altura debería ser la de la lámina de plasma, es decir, de 10 a 12 radios terrestres. Esta curiosa estructura magnética, con el plasma caliente que confina, constituye un plasmoide. Una vez se ha formado, las líneas de campo abiertas y circundantes de los lóbulos se reconectan, formando una vaina de líneas de campo interplanetarias que actúan como una honda e impulsan el plasmoide hacia la cola a velocidades de 500 a 1000 kilómetros por segundo, es decir, de 5 a 10 radios terrestres por minuto. Entre la línea neutra y el plasmoide en marcha, sólo queda una capa de plasma muy delgada entretejida de líneas de campo interplanetarias.




**6. CREACION DE UN PLASMOIDE: uno de los procesos a través de los cuales la cola del campo magnético libera energía. El proceso, predicho por la teoría, se confirmó gracias a los datos recogidos por satélites en 1983. El primer dibujo *(a)* muestra la cola del campo magnético una hora antes del comienzo de una subtormenta magnética. La línea neutra distante se encuentra a unos 100 radios terrestres corriente abajo de la Tierra. La subtormenta se desencadena cuando la cesión de energía magnética del viento solar a la magnetosfera alarga en exceso las líneas de campo de la cola del campo magnético y forma una línea neutra propia de la subtormenta dentro de la capa de plasma *(en color)*, a unos 15 radios terrestres corriente abajo de la Tierra *(b)*. Las líneas de fuerza se reconectan rápidamente en la línea neutra de la subtormenta *(c)* hasta que un plasmoide –plasma caliente aprisionado en un conjunto concéntrico de líneas de fuerza– queda libre de ligaduras magnéticas con la Tierra *(d)*. Entonces, las líneas de los lóbulos *(en gris)* comienzan a reconectarse, formando una vaina de líneas de fuerza interplanetarias que actúa como una honda, impulsando el plasmoide corriente abajo *(e)*. Las diversas reconexiones forman también líneas cerradas que se contraen hacia la Tierra, inyectando partículas de alta energía en los cinturones de radiación de Van Allen y depositando en la alta atmósfera la energía que produce las auroras asociadas a las subtormentas. La línea neutra de la subtormenta magnética se mueve corriente abajo *(f)*, restableciendo la configuración normal de la cola.**

Hasta ese momento no han transcurrido más que unos 10 minutos desde el comienzo de la subtormenta. Sigue después un intervalo de entre 30 minutos y una o dos horas, durante el cual la línea neutra continúa en su posición cercana a la Tierra y sigue siendo sede de reconexión de líneas de campo magnético, formando una familia de líneas cerradas del lado de la Tierra dirigido hacia la cola y originando nuevas y brillantes auroras. El intervalo acaba con la brusca y rápida retirada de la línea neutra hacia el espacio. El plasma proyectado hacia la Tierra desde la línea neutra en retirada rellena la capa de plasma, restableciendo su configuración anterior a la subtormenta.

El modelo de línea neutra de los sucesos magnetosféricos asociados a una subtormenta recibió pronto apoyo de las observaciones desde satélite de los campos magnéticos, los plasmas y las partículas individuales de alta energía en la región de la cola magnética que se extiende de 7 a 35 radios terrestres corriente abajo desde la Tierra. De hecho, el modelo se desarrolló, parcialmente, como interpretación de esas observaciones. Entonces, en 1983, el satélite *Explorador Internacional Sol-Tierra 3, ISEE-3,* lanzado por la Administración Nacional de Aeronáutica y del Espacio (NASA) realizó cortes de la cola del campo magnético a distancias de hasta 230 radios terrestres. Los datos que envió a la Tierra confirmaron de una manera espectacular el modelo de línea neutra: revelaron el paso de los plasmoides predichos por el modelo.

El estudio de la cola del campo magnético terrestre tiene implicaciones más allá del interés por explicar los fenómenos electromagnéticos en el espacio. Unica entre los reinos de plasma en su accesibilidad mediante satélites, la cola del campo magnético terrestre constituye, en efecto, un laboratorio donde los procesos físicos en los plasmas astronómicos se pueden estudiar con un detalle inigualado. La reconexión magnética es uno de esos procesos; se trata, quizá, del más importante en muchos fenómenos astronómicos. Otro ejemplo son los chorros astrofísicos: chorros de plasma caliente largos y estrechamente colimados que se extienden desde las estrellas, e incluso desde las galaxias. Se sabe todavía poco acerca de su origen o su dinámica. Pero sí se conoce en ellos la presencia de campos magnéticos; algunos chorros contienen subestructuras coherentes cuyo aspecto sugiere que el plasma de su interior está confinado magnéticamente. Tales subestructuras son, así cabe pensarlo, variantes cosmológicas de los plasmoides.

Las colas de campos magnéticos se dan frecuentemente en el sistema solar. Se forman dondequiera que un cuerpo que tiene un campo magnético intrínseco (como en el caso de Mercurio, la Tierra, Júpiter y Saturno) o un cuerpo que tiene una atmósfera eléctricamente conductora (caso de Venus y los cometas) está inmerso en un plasma magnetizado en movimiento. Bajo esas condiciones, el flujo de plasma junto al cuerpo queda entorpecido por las fuerzas electromagnéticas aquí descritas, almacenándose energía en forma de líneas de campo magnético extendidas en el espacio corriente abajo del cuerpo. Es probable que la liberación intermitente de energía se logre mediante reconexión magnética, de manera parecida a como sucede en la magnetosfera terrestre. De hecho, se ha ofrecido esta explicación para los sucesos de desconexión, o desprendimiento, de porciones de las colas de los cometas, que aparecen frecuentemente a la observación.

Un paso ulterior en el estudio de la cola del campo magnético terrestre, paso en el cual el autor toma parte, es el PROMIS, o Estudio Internacional de la Región Polar y la Magnetosfera Exterior. Iniciado a mediados de marzo, continuará hasta mediados de junio. La idea es coordinar la actividad de satélites europeos y estadounidenses ya en el espacio de modo que se puedan relacionar las imágenes de auroras obtenidas simultáneamente en las regiones polares Norte y Sur de la Tierra con los datos recogidos, simultáneamente, en la cola del campo magnético, distante de la Tierra.

Finalmente, debe observarse que la reconexión magnética y la formación de plasmoides se produce en los plasmas de laboratorio de igual manera que en el espacio. Como parte del esfuerzo internacional para dominar la energía de la fusión nuclear, se están desarrollando dispositivos de prueba en los que un plasma muy caliente se mantiene confinado por campos magnéticos mientras sus núcleos atómicos constituyentes interaccionan. En algunos experimentos recientes, las líneas del campo magnético que pasan por un cuerpo de plasma son arrancadas de su fuente externa por reconexión magnética, dejando un plasma confinado en bucles magnéticos cerrados. El resultado es un plasmoide muy parecido a los que se producen en la cola del campo magnético terrestre.

# Genética molecular de la hemofilia

*Se desangran los hemofílicos porque un defecto génico les priva de una proteína fundamental en la coagulación de la sangre. Por aislamiento del gen normal y su cultivo en células se ha logrado ahora la síntesis artificial de ese factor*

Richard M. Lawn y Gordon A. Vehar

Un minúsculo defecto en un gen humano, y el consiguiente déficit o carencia de la proteína que éste determina, puede ser causa de una grave enfermedad. Es el caso de la hemofilia. La sangre de los hemofílicos carece de una proteína a la que corresponde un papel crucial en la cascada de reacciones enzimáticas que conducen a la coagulación de la sangre dondequiera que se produzca una herida. De no recibir tratamiento, la más leve contusión provoca hemorragias internas en los hemofílicos graves, que suelen morir a temprana edad tras la crisis derivada de alguna herida.

Afortunadamente, cabe administrarles transfusiones regulares de un concentrado de la proteína que les falta. Desde principios de los años sesenta, cuando se inició ese tipo de tratamiento, la vida de los hemofílicos ha mejorado espectacularmente en los países desarrollados y, su esperanza de vida, de sólo 20 años hasta hace poco, es ahora casi la normal. No obstante, para preparar el concentrado proteínico se necesita la sangre de muchos donantes, y ello encarece el tratamiento. En Occidente, la cantidad necesaria para el tratamiento medio de un solo hemofílico cuesta entre 900.000 y 1.500.000 pesetas al año. En los países pobres a menudo resulta imposible disponer del concentrado. Además, al proceder la sangre de personas muy distintas, se corre el riesgo de propagar enfermedades víricas. La mayoría de los hemofílicos sufren infección crónica del virus de la hepatitis y corren hoy el riesgo de contraer el síndrome de inmunodeficiencia adquirida (SIDA).

No sorprendió, por tanto, el interés puesto en elaborar la proteína anthimofílica por medio de técnicas de ingeniería genética. En la mayoría de los casos la hemofilia se debe a un defecto localizado en el segmento de ADN que determina la proteína denominada factor VIII, necesaria para la coagulación. Investigadores de dos compañías de biotecnología, entre ellos nuestro grupo de Genentech Inc., de San Francisco, han aislado recientemente el gen del factor VIII a partir de células de individuos normales y lo han recombinado con el ADN de células de cultivo. Las células así recombinadas se replican, es decir, producen muchas copias del gen del factor VIII. En todas las células recombinadas se expresan las instrucciones contenidas en el gen, sintetizándose en conjunto una cantidad considerable de factor VIII.

La proteína obtenida por ingeniería genética es funcional. En pruebas de laboratorio coagula la sangre extraída de hemofílicos, y se ha comprobado su eficacia en perros hemofílicos. Tanto Genentech como el Instituto de Genética de Cambridge, la otra empresa implicada en estas investigaciones, se afanan en poner a punto métodos para la síntesis del factor VIII y su consiguiente explotación comercial. Si bien la proteína debe superar aún nuevas pruebas en animales y posteriormente en humanos, parece probable que en los próximos años se cuente en el mercado con cantidades suficientes de factor VIII puro y libre de virus.

La perspectiva de disponer del gen clonado está ya transformando los estudios sobre la hemofilia, dificultados por la extraordinaria complejidad que entraña purificar el factor VIII a partir de la sangre. Además de escaso, el factor VIII es una proteína larga e inestable. Cuando empezamos nuestros trabajos se desconocía su estructura y el lugar de su síntesis. Pudiéndose elaborar de forma artificial, su estudio resultará mucho más fácil. De hecho, se ha deducido ya la estructura fundamental de la proteína a partir de la secuencia del gen. En algunos casos, incluso, se han precisado las mutaciones que determinan la hemofilia y que se transmiten de una generación a otra.

El conocimiento del carácter hereditario de la hemofilia se remonta a los autores del Talmud: decretaron que los niños cuyos hermanos mayores o primos se desangraran hasta morir tras la circuncisión no tenían obligación de someterse a ella. El patrón característico de herencia de la enfermedad (sólo suele afectar a los varones, aunque las hembras pueden ser portadoras) se dilucidó a principios del siglo XIX. Quizá la portadora más famosa fue la reina Victoria de Inglaterra. Uno de sus hijos era hemofílico, y portadoras al menos dos de sus hijas . Con los matrimonios de éstas, el gen mutante de la reina Victoria se propagó entre las familias reales de Alemania, Rusia y España.

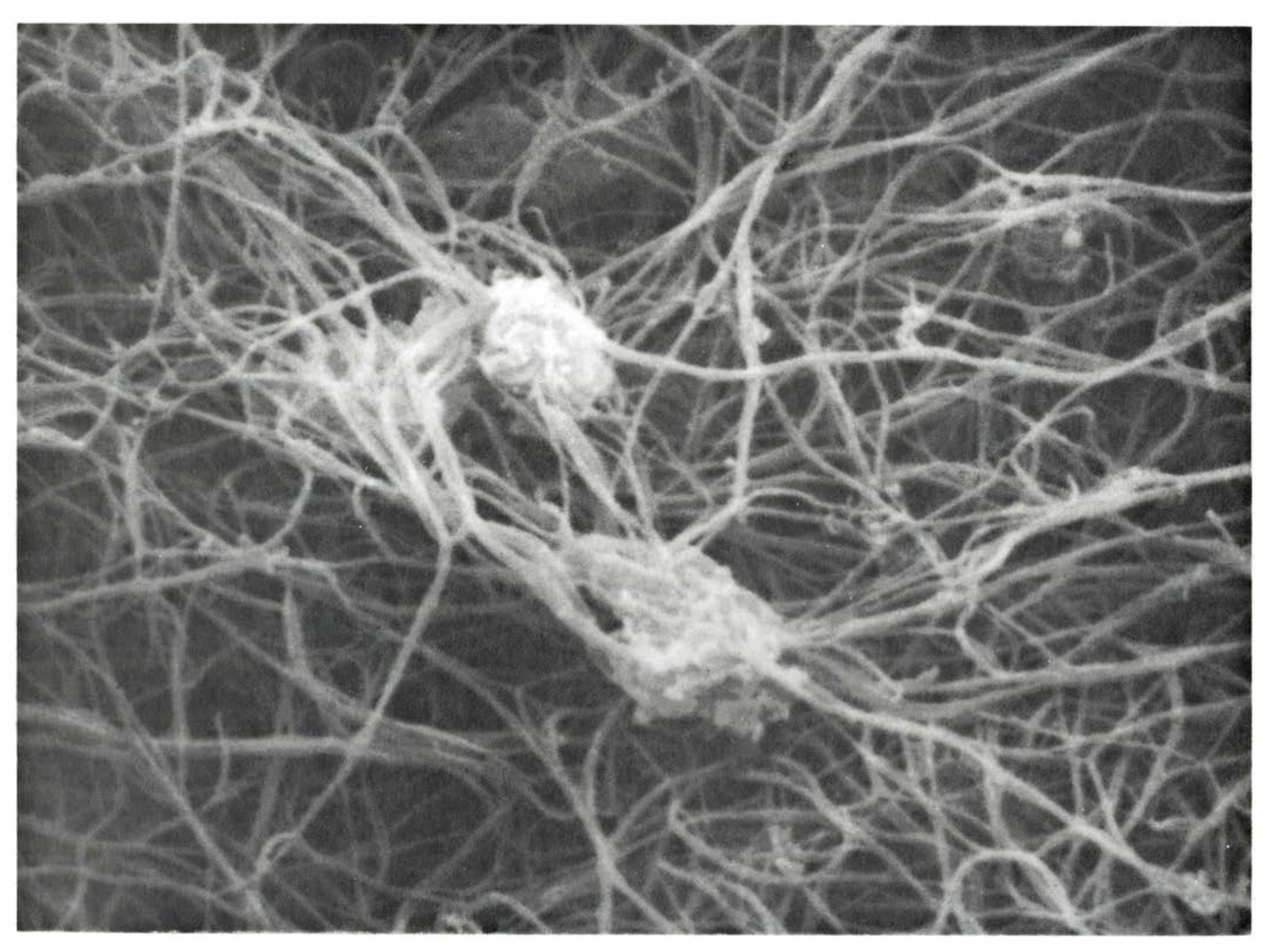
Sabemos hoy que la hemofilia es una enfermedad ligada al sexo, pues el gen que determina el factor VIII se encuentra en el cromosoma X. Las células de las mujeres contienen dos cromosomas X, mientras que las de los varones contienen un cromosoma X y otro Y. El varón, al tener sólo un gen para el factor VIII, heredado de su madre, será hemofílico si dicho gen es defectuoso. Una hembra, por el contrario, tiene dos genes para el factor VIII, uno heredado de la madre y otro del padre. Puede, por tanto, llevar un gen defectuoso sin sufrir la enfermedad, ya que el gen normal del otro cromosoma X le protege. Sólo en raras ocasiones los dos genes serán defectuosos, por lo que habrá pocas hembras hemofílicas. Por termino medio, las mujeres portadoras pasarán su gen mutante a la mitad de los hijos, que serán hemofílicos.

Dominamos sólo las líneas generales del proceso de la coagulación sanguínea, que falla en los hemofílicos. Se inicia en las plaquetas, que se adhieren al lugar donde se haya producido una herida. Las plaquetas se desprenderían fácilmente de no ser por la formación inmediata de cadenas de fibrina, un polímero insoluble, que las fijan en su sitio. La genésis de una malla de fibrina a partir de su precursor soluble, el fibrinógeno, constituye el fenómeno decisivo de la coagulación. Es el resultado final de una compleja cascada de interacciones que se desencadenan al sufrir agresión un vaso sanguíneo. En cada paso de tal cascada un precursor proteico sufre un corte específico y se transforma en una enzima activa, una proteasa. A continuación, la proteasa escinde otra proteína, convirtiéndola en una proteasa distinta. La mayoría de esas etapas necesitan cofactores; en algunos casos son también proteínas, que suelen presentar dos estados: una forma activa y otra inactiva. El factor VIII, pese a su nombre, constituye uno de esos cofactores. A mitad de la cascada colabora con el factor IX en la activación del factor X.

En el proceso de la coagulación intervienen circuitos de retroalimentación positiva para acelerar la respuesta y otros de signo contrario para detenerla. Así, por ejemplo, la trombina, proteasa que convierte el fibrinógeno en fibrina, activa el factor VIII; al propio tiempo, activa, sin embargo, la proteína C, proteasa también, que activa el factor VIII. La concentración de factor VIII en la sangre normal es tan baja (por cada molécula de factor VIII hay aproximadamente un millón de moléculas de albúmina, el principal componente proteico de la sangre) que puede considerarse un factor limitante. En otras palabras, la rápida activación y desactivación del factor VIII explica, en parte, el delicado equilibrio existente en los individuos normales entre la formación del coágulo y el libre flujo de la sangre.

En los hemofílicos, ese balance se rompe. Un 85 por ciento de esos enfermos, aproximadamente un varón de cada 10.000, padece la hemofolia clásica (la del tipo *A*); su carencia de factor VIII funcional detiene la cascada de la coagulación antes de que se pueda formar la fibrina. El resto pacede, casi en su totalidad, hemofilia *B*, resultante del déficit de factor IX. Se ha clonado ya el factor IX, y varias empresas están intentando obtenerlo por aplicación de técnicas de ingeniería genética. El producto más ansiado, no obstante, es el factor VIII, pues la incidencia de la hemofilia *A* supera con mucho a la de la hemofilia *B*.

La fabricación de una proteína del tamaño y la escasez del factor VIII ha supuesto un desafío tecnológico sin precedentes. Las dificultades nos han obligado a modificar los métodos empleados habitualmente en la clonación y manipulación de genes para que fabriquen proteínas.



**1. CADENAS DE FIBRINA que estabilizan un coágulo sanguíneo en el sitio donde se ha abierto una herida. Se enredan en ellas gran cantidad de plaquetas, que constituyen el grueso del coágulo. La micrografía, realizada por Jon C. Lewis, muestra un coágulo formado en una suspensión de plaquetas y fibrina. En la sangre, el coágulo es el producto final de una compleja cascada de reacciones enzimáticas que culminan con la conversión del fibrinógeno, una proteína soluble, en las cadenas insolubles de fibrina. En los hemofílicos, falta, o es defectuosa, una proteína esencial en la cascada que lleva a la coagulación.**

El factor VIII, como toda proteína, está constituido por aminoácidos. El gen que lo determina es un segmento de ADN, una cadena de nucleótidos. La secuencia de los aminoácidos viene determinada por la de los nucleótidos, cada uno de los cuales lleva una base: adenina (*A*), timina (*T*), guanina (*G*) o citosina (*C*). Cada tres bases, o codón, determinan un aminoácido. Las bases presentan relaciones de complementariedad: la adenina forma puentes de hidrógeno con la timina y, la guanina, con la citosina. Los apareamientos entre bases mantienen la estructura en doble hélice de las dos cadenas de ADN. Además, condicionan la transcripción de un gen en ARN mensajero (ARNm) y la posterior traducción de éste a proteína.

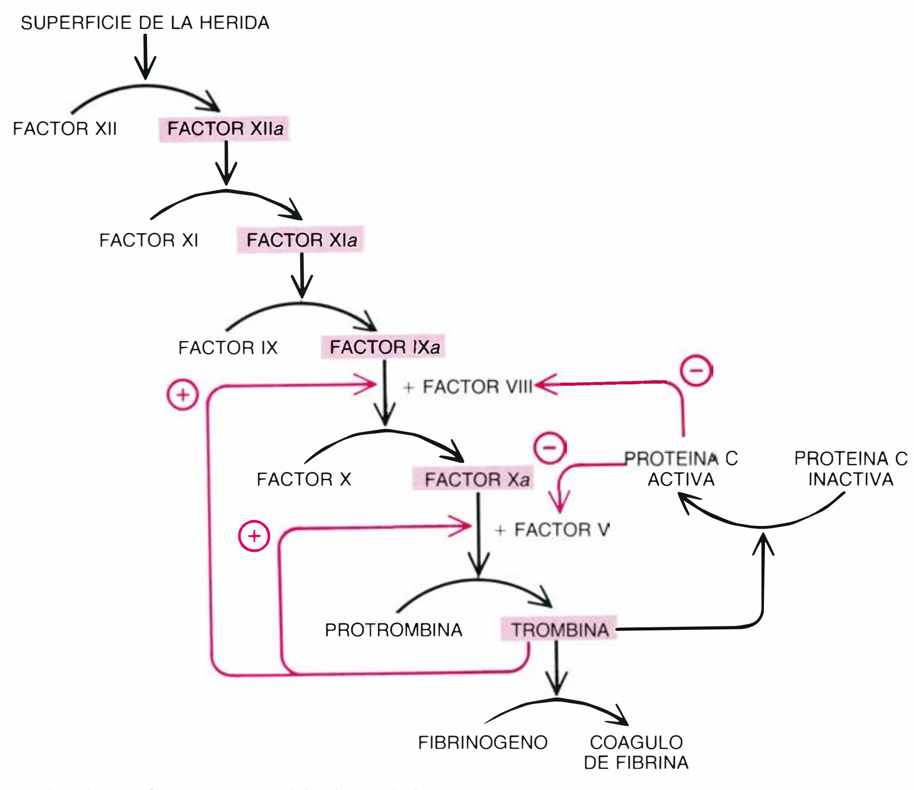
A la hora de fabricar una proteína en el laboratorio el principal problema consiste en encontrar, entre los miles que posee cualquier célula, el gen que la determina. La complementariedad entre las bases permite superar ese obstáculo. Para localizar el gen basta con un pequeño segmento de ADN o ARN cuya secuencia de bases complemente parte de aquél; ese breve segmento hará las veces de sonda. Se obtiene una sonda de ADN, por ejemplo, retrotranscribiendo, de acuerdo con las reglas del código genético, parte de la secuencia de aminoácidos de la proteína deseada. Se "marca" con un nucleótido radiactivo y, al mezclarla con una "genoteca" (la "biblioteca" de genes de un organismo), hibridará sólo con el gen desado, que así quedará también marcado.

Cuanto menor sea la genoteca, más fácil resultará seleccionar un gen específico. En el protocolo habitual de clonación, denoninado clonación por ADNc, se reduce el tamaño de la genoteca aprovechándose de la circunstancia de que no todos los genes actúan en todas las células. En una célula determinada, sólo parte de los genes se transcriben a ARNm y posteriormente se traducen a proteína. Si se sabe qué tipo celular fabrica la proteína deseada, basta con rebuscar entre las moléculas de ARNm de esas células. Entre ellas debe encontrarse el producto de la transcripción del gen deseado.

Para encontrar el gen es preciso retrotranscribir todos los ARNm en ADN, lo que se hace con la ayuda de la enzima denominada transcriptasa inversa. Los segmentos de ADN así copiados, el ADNc, se unen enzimáticamente a un vehículo genético, un vector, por ejemplo el bacteriófago lambda. Se introducen los fagos en bacterias, de tal forma que se multiplique cada uno de ellos en una región distinta de una cápsula de petri. Al multiplicarse, los fagos matan las bacterias y forman un halo de lisis en su derredor. El conjunto de los halos constituye una genoteca. Al menos uno de ellos debe contener el fragmento de ADNc deseado. El halo en cuestión se identifica por hibridación con una sonda.

La estrategia de clonación con ADNc sólo funciona si se sabe qué células del cuerpo elaboran la proteína deseada. El éxito será mayor si, además, la proteína se produce en abundancia, ya que en ese caso las células contendrán muchas copias del ARNm y, por tanto, muchos serán también los halos de la genoteca de ADNc que constengan copias del gen. Nada de eso resulta aplicable al factor VIII. Escasea la proteína y, al dar comienzo nuestros trabajos, nadie sabía qué órgano la producía. De habernos decidido a formar una genoteca de ADNc nos hubiéramos arriesgado a elegir un tipo equivocado de células y obtener una genoteca que no contuviese el gen para el factor VIII.

Decidimos, por tanto, buscar el gen del factor VIII donde con seguridad lo encontraríamos: en una genoteca construida a partir del genoma entero, la dotación completa de genes de la célula. Para formar una genoteca genómica se extraen los cromosomas de las células, se corta el ADN en fragmentos con ayuda de enzimas específicas y se unen éstos al ADN del fago lambda. La genoteca genómica contiene cientos de veces más ADN que la de ADNc, por lo que se complica la búsqueda con la sonda.




**2. CASCADA QUE LLEVA A COAGULACION. Se inicia cuando las células dañadas activan el denominado factor XII. Termina con la conversión del fibrinógeno en fibrina, por mediación de la trombina. En cada paso, una proteína inactiva se transforma en proteasa, enzima capaz de escindir otra proteína (*color*), que a su vez activa el siguiente polipéptido. Algunos pasos requieren el concurso de cofactores, así los factores VIII y V. En la cascada se generan circuitos de retroalimentación positiva y negativa (*flechas de color*). La trombina activa los factores VIII y V, pero también los inactiva por mediación de la proteína *C*. Ello permite mantener bajo control el proceso de la coagulación. El 85 por ciento de los hemofílicos carece de factor VIII.**

Antes de fabricar cualquier sonda debíamos conocer parte de la secuencia de aminoácidos del factor VIII. No resultó ello tarea fácil. La proteína no se purificó hasta 1980, cuando uno de nosotros (Vehar), a la sazón en el laboratorio de Earl W. Davie, de la Facultad de Medicina de la Universidad de Washington, laboriosamente extrajo varios miligramos de factor VIII puro a partir de 25.000 litros de sangre de vaca. Posteriormente, Edward Tuddenham y sus colegas, del Royal Free Hospital de Londres, obtuvieron suficiente factor VIII humano para que los investigadores de Genentech secuenciaran un pequeño segmento de la proteína. Un grupo del Instituto de Genética obtuvo resultados similares con factor VIII de cerdo, purificado por David N. Fass, de la Clínica Mayo.

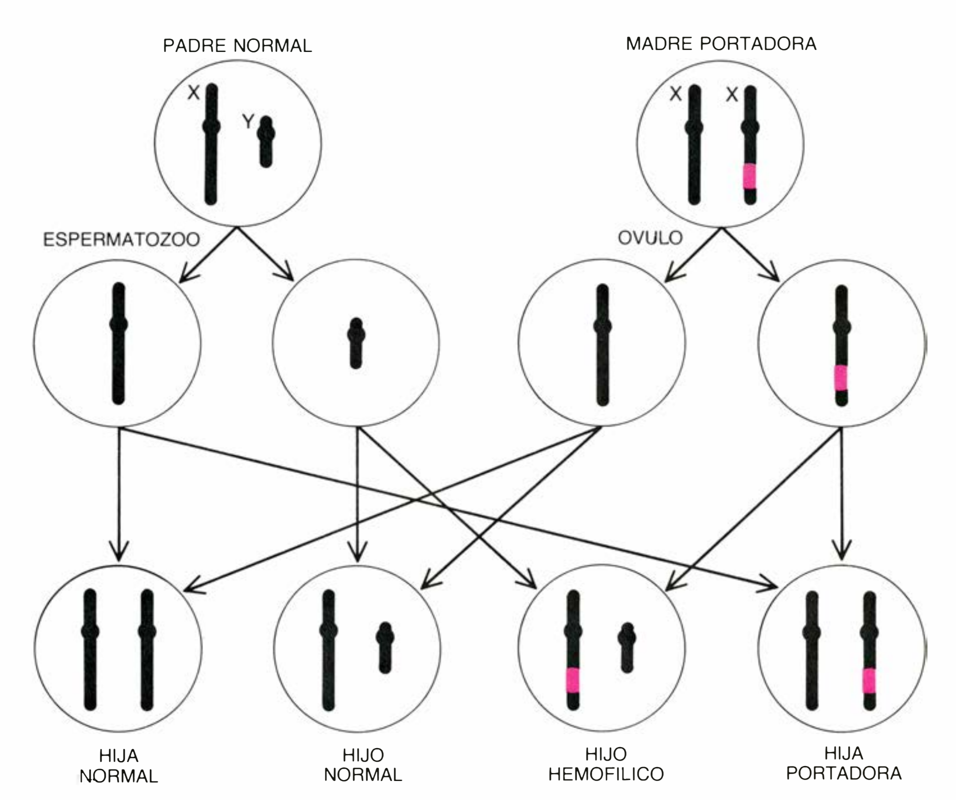
El siguiente paso fue obtener la se-

cuencia del gen a partir de la proteína. Debe, al efecto, afrontarse un nuevo problema: el código genético es redundante. Ciertos aminoácidos vienen determinados hasta por seis codones diferentes. (Existen 64 posibles codones de tres bases, y sólo 20 aminoácidos diferentes.) Se obvia el problema sintetizando un conjunto de sondas de ADN de pequeño tamaño (unas 17 bases) que cubran todas las posibilidades. Cuanto más corta sea la sonda, sin embargo, mayor es la probabilidad de que complemente aleatoriamente una secuencia de ADN que no sea la del gen deseado. Cuando se utilizan muchas sondas pequeñas para escudriñar una gran genoteca genómica, llega a resultar problemática la aparición de falsos positivos.

Para evitarlo, nos servimos de una sola sonda, relativamente grande (36 bases), obtenida a partir de una secuencia de 12 aminoácidos del factor VIII. Hubo que elegir entre las 147.456 posibles formas de componer dicha secuencia. Afortunadamente, sabíamos que en los genes de mamíferos algunos codones son más frecuentes que otros. Nuestras conjeturas no iban desencaminadas; acertamos 30 de las 36 bases correctas, lo que venía a ser casi una concordancia absoluta. Al efectuar la búsqueda en la genoteca, la sonda sintética hibridó con fragmentos del ADN del factor VIII que se solapaban: habíamos identificado halos que contenían diferentes partes del gen.

El gen completo es tan grande que no cabe en un solo fago. Para dar con el gen entero, William I. Wood, Jane Gitschier y otros investigadores, de nuestro laboratorio, volvieron a buscar en la genoteca, pero ahora utilizaron como sonda los fragmentos del propio gen. Por iteración del procedimiento, técnica ésta denominada "paseo por el cromosoma", obtuvieron por fin una serie de segmentos que se solapaban y que abarcaban la secuencia entera del gen.

Mide éste unas 186.000 bases. Sin embargo, la información que determina el factor VIII se encuentra dispersa entre 26 exones (secuencias informativas) que, juntas, vienen a suponer menos del 5 por ciento de la longitud total del gen. Ello se debe a que los exones están separados por 25 intrones, o secuencias espaciadoras, no informativas. Una vez transcrito en ARN el gen completo, los intrones se eliminan y los exones se empalman constituyendo el ARNm, que dirige la síntesis de proteínas. Para fabricar factor VIII en un cultivo celular, también precisábamos nosotros un gen sin intrones: debíamos encontrar el ARNm del factor VIII y convertirlo en ADNc.

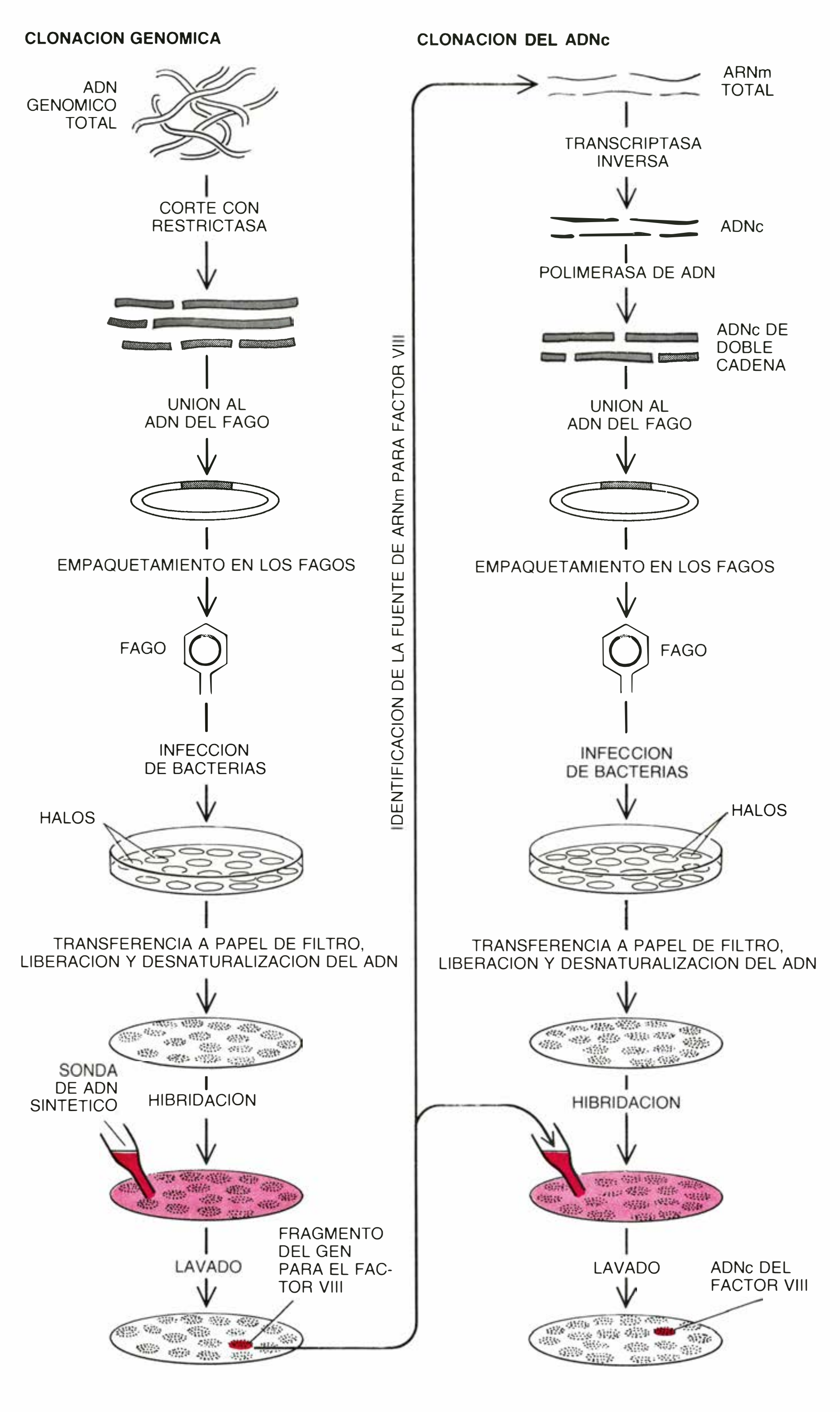



**3. HERENCIA DE LA HEMOFILIA, que está ligada al sexo porque el gen del factor VIII se halla en el cromosoma X. Un varón portador de un gen del factor VIII que haya mutado carece de factor VIII normal; es, por tanto, hemofílico. Una hembra portadora no sufre la enfermedad, pues el gen normal de su segundo cromosoma X la protege; la mitad de sus hijas serán portadoras y la mitad de sus hijos hemofílicos. En el caso de que el padre fuese hemofílico y la madre normal (no contemplado en el esquema), los hijos varones no serían hemofílicos, puesto que reciben del padre el cromosoma Y, pero todas las hijas serían portadoras.**

Disponiendo ya de sondas constituidas por fragmentos del auténtico gen era tarea sencilla encontrar qué células fabricaban factor VIII y su mensajero. Si la sonda no hibridaba con los ARNm de algún tipo celular, podíamos ahora estar seguros de que no era problema de la sonda, confianza que no cabía con la sonda sintética. Un grupo dirigido por Daniel Capon, de Genentech, encontró señales positivas de hibridación entre los mensajeros de una línea celular cultivada y, John J. Toole y sus colegas, del Instituto de Genética, comprobaron que las células del hígado constituían una fuente de tales mensajeros. Aunque el hígado parece ser la fuente principal, se ha encontrado también ARNm del factor VIII en otros tejidos, así riñones, bazo y células linfáticas. En individuos normales, la mayor parte del factor VIII probablemente se sintetice en el hígado y se segregue luego al torrente sanguíneo.

Descubierta por fin una fuente de ARNm del factor VIII, los investigadores de ambas compañías formaron genotecas de ADNc. De nuevo utilizaron, por sondas, fragmentos del gen clonado, esta vez para "pescar" el ADNc del factor VIII en una genoteca de miles de clones. En realidad, hubo que ensamblar ese ADNc a partir de varios segmentos solapantes, pues el ARNm tenía unas 9000 bases y las técnicas actuales no permiten copiar de una vez moléculas tan largas. Finalmente debió añadirse al ADNc las secuencias de control, que indican a las enzimas correspondientes de la célula recombinante dónde deben iniciar y concluir la transcripción de un gen.

Las bacterias recombinantes, por lo general *Escherichia coli*, resultaban idóneas para elaborar las primeras proteínas producidas por ingeniería genética (insulina e interferón), pues se trata de moléculas baastante pequeñas. Sin embargo, no suelen estar convenientemente equipadas para producir una proteína grande y compleja como el factor VIII, ni suelen tener las enzimas pertinentes para modificar o plegar una gran proteína después de su síntesis. Decidimos, por esa razón, insertar en células de hámster, de fácil cultivo en el laboratorio, el gen del factor VIII clonado. Las células recombinantes del hámster secretaron factor VIII al medio de cultivo. Mejor dicho, esperábamos que fuese factor VIII. No podíamos estar seguros de la actividad



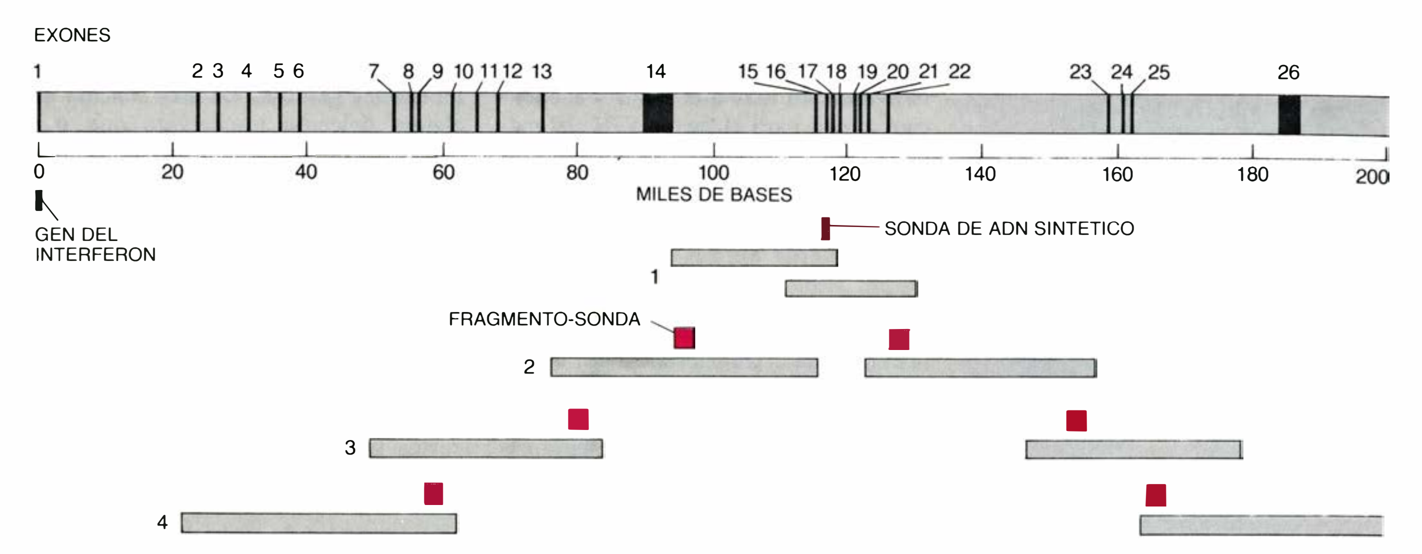

**4. CLONACION DE GENES. Obliga a descubrir una secuencia específica entre las miles de que consta una célula humana. El método habitual, cuando se conoce la célula que elabora la proteína deseada, es escudriñar una genoteca de copias de ADN (ADNc) obtenidas por retrotranscripción de los ARN mensajeros (ARNm) de dichas células (*derecha*). Cuando se buscó el gen del factor VIII, sin embargo, los autores desconocían dónde se producía la proteína, por lo que hubo que analizar todo el genoma humano (*izquierda*). Al efecto se introdujeron fragmentos de ADN cromosómico en el ADN del bacteriófago lambda. Se multipicaron los fagos, cada uno de los cuales portaba un fragmento de ADN humano, formando halos de lisis sobre un cultivo bacteriano. Para identificar el halo que contenía el gen del factor VIII se transfirieron los fagos a un papel de filtro y se les provocó la liberación del ADN. Expuesto éste a una sonda radiactiva (un fragmento de ADN sintético correspondiente a parte del factor VIII), hibridó con parte del gen del factor VIII, lo que permitió su aislamiento. Para producir factor VIII en células cultivadas debió, además, fabricarse ADNc de ese factor a fin de eliminar los intrones (secuencias no informativas) del gen. Para ello pueden utilizarse, a modo sonda, los fragmentos del gen clonados previamente, primero para identificar las células que fabrican ARNm del factor VIII y, posteriormente, para identificar el ADNc correspondiente en la genoteca entera de ADNc.**

de la proteína hasta comprobar que coagulaba normalmente la sangre de los hemofílicos. Podía darse el caso, por ejemplo, de que nuestra proteína constituyera sólo una parte del complejo proteínico que les falta a los hemofílicos, o incluso que, dadas las dificultades registradas en la purificación del factor VIII, hubiésemos clonado un gen que determinase alguna impureza de la preparación. Había que considerar todas esas inquietantes posibilidades, aunque a la postre ninguna resultó ser cierta. El factor VIII obtenido por ingeniería genética coagulaba la sangre de hemofílicos; es más, su comportamiento, así se ha comprobado, es equivalente en todo sentido al de la proteína sanguínea.

¿Qué estructura tiene el factor VIII? La disponibilidad de las modernas técnicas de clonación de genes ha dado lugar a una situación insólita, cual es la de determinar indirectamente la secuencia de algunas proteínas a partir de la secuencia de ADN del gen. Es el caso del factor VIII, polipéptido demasiado largo y escaso para secuenciarlo entero de forma directa. Antes de clonar el gen, los investigadores discrepaban incluso respecto de su tamaño aproximado. Las estimaciones variaban hasta en un factor cercano a 100.

Puede ya darse respuesta a esa incógnita. Los 9000 nucleótidos del ADNc del factor VIII determinan una proteína de una longitud de 2351 aminoácidos. (Unas 2000 bases situadas en los extremos del gen se transcriben en ARNm, pero no se traducen a proteína.) Los primeros 19 aminoácidos forman una secuencia hidrofóbica, típica de las proteínas que se secretan. Ese "péptido de señalización" suele eliminarse de la proteína durante su secreción; así pues, la molécula madura de factor VIII debe estar formada por los restantes 2332 aminoácidos. Su peso molecular debe situarse en torno a los 330.000 dalton (un dalton equivale a $1,66 \times 10^{-24}$ gramos), en tanto una molécula de interferón, por ejemplo, no consta más que de 166 aminoácidos y pesa unos 19.000 dalton. El gen del factor VIII es, con mucho, el más largo de los que se haya clonado y expresado en una célula foránea.

El análisis de la secuencia de aminoácidos del factor VIII revela que la proteína consta de segmentos similares repetidos, tres de los cuales se han denominado *A*. De los aproximadamente 350 aminoácidos que componen la se-




**5. EL GRAN TAMAÑO del gen del factor VIII, el mayor gen clonado hasta la fecha, obligó a los investigadores a utilizar la técnica de clonación denominada "paseo por el cromosoma". Contiene el gen unas 186.000 bases, mientras que el del interferón, por ejemplo, clonado en1980, no posee más que unas 600. Tamaña longitud impide que quepa en un solo fago, de ahí que tuviera que recuperarse el gen del factor VIII a partir de fragmentos ubicados en diversos halos de la genoteca genómica. La sonda de ADN sintético utilizada para buscarlo hibridó con fragmentos que se solapaban (*1*). A su vez, partes de esos segmentos se utilizaron de sonda en la búsqueda de más fragmentos por la genoteca (*2*). Repitiendo varias veces el procedimiento se identificó prácticamente todo el gen (*3, 4*). Los exones, o secuencias informativas (*bandas negras*), representan menos del 5 por ciento del gen. Los 26 exones están separados por 25 intrones.**

cuencia de cada segmento *A*, aproximadamente una tercera parte son comunes a los otros tres segmentos. Un nivel comparable de homología se aprecia también entre los dos segmentos de 150 aminoácidos que se denominan *C*, que no guardan homología con los *A*. Dada la existencia de 20 aminoácidos diferentes, la homología entre un tercio de las secuencias casi con seguridad no constituye un hecho aleatorio. Los segmentos homólogos deben estar emparentados.

La comparación, por ordenador, de las secuencias del factor VIII con las de otras proteínas reveló un dato sorprendente acerca de la historia evolutiva de dicho factor. Los tres segmentos *A* tienen un grado de similitud con los tres dominios de la ceruloplasmina (una proteína sanguínea de transporte de cobre) comparable al que tienen entre sí. Aunque a priori no había razón para sospechar la existencia de ningún parentesco entre el factor VIII y la ceruloplasmina, la homología entre sus dominios sugiere que sus genes evolucionaron a partir de un antecesor común. La proteína ancestral quizá constara de tres dominios idénticos, precursores de los segmentos *A*. En la ceruloplasmina, los tres dominios *A* aparecen contiguos y componen la molécula entera, mientras que están separados por casi 1000 aminoácidos. En el gen del factor VIII esa región intermedia está determinada por un solo exón, enorme, que pudo insertarse en el gen ancestral. De igual forma, las regiones que determinan los segmentos *C* del factor VIII se añadieron al extremo del gen.

El factor VIII purificado a partir de sangre de donantes casi nunca resulta idéntico a la molécula completa de 330.000 dalton determinada por el gen. Se supone que el factor VIII sufre una serie de cortes al activarse en la sangre. Uno de esos cortes elimina la región situada entre el segundo y tercer dominios *A* (la región *B*). Ello es lo que se deduce cuando se comparan las secuencias de pequeños fragmentos de la proteína activa con la secuencia del gen clonado. El factor VIII activo parece constar de una subunidad de 90.000 dalton unida a otra de 73.000. La primera está formada por los dos primeros dominios *A*. El péptido que los une se elimina. La segunda subunidad está formada por el tercer dominio *A* y los dos dominios *C*. De la misma manera que esos cortes específicos activan la molécula, otros, producidos dentro de las subunidades, la inactivan inmediatamente. Probablemente interrumpan esas reacciones la cascada de la coagulación sanguínea en el momento oportuno.

Distan de estar claros los detalles del funcionamiento del factor VIII en la cascada. Se encuentra unido a una proteína de transporte denominada factor de von Willebrand, que mantiene al factor VIII en la circulación sanguínea y quizá le ayude a situarse en la superficie de una plaqueta donde se haya producido una herida. Ya sobre la plaqueta, el factor VIII se separa del de von Willebrand y forma un complejo con los factores IX y X. Se ignora de momento cuáles son los sitios de unión a esas proteínas. Todo lo que sabemos es que, en ausencia del factor VIII, el factor IX no logra activar al X.

En el estudio del modo de acción de una proteína resulta habitual analizar algunas de sus formas anormales obtenidas por mutación. Estamos empezando a aplicar esa estrategia al factor VIII, lo que simultáneamente nos está permitiendo identificar las mutaciones que causan la hemofilia. Todo ello habrá de traducirse en una mejora del conocimiento de la enfermedad y de su terapéutica.

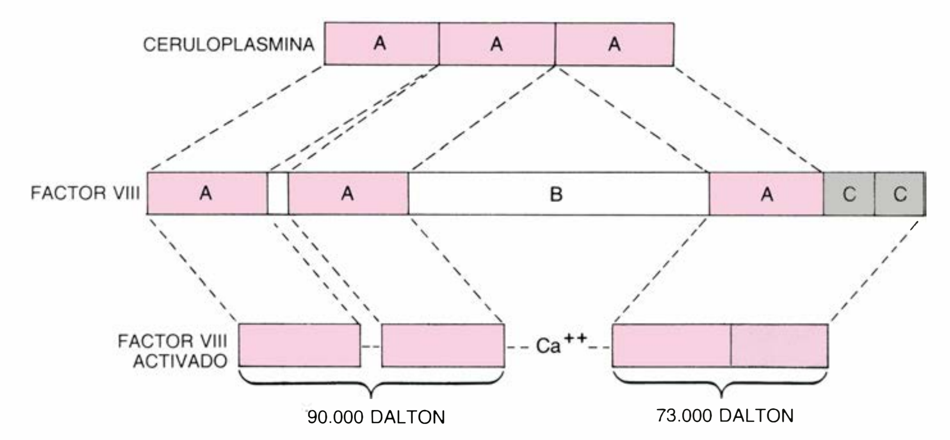
Como era de prever, se ha encontrado que la base molecular de la hemofilia es heterogénea. Hace 50 años, el genetista J. B. S. Haldane señaló que las enfermedades graves ligadas al cromosoma X debían de originarse constantemente por efecto de mutaciones espontáneas aleatorias, ya que, de no ser así, tales enfermedades acabarían por desaparecer. De hecho, casi un tercio de los casos de hemofilia que se observan actualmente ocurren en familias que carecen de historial en cuanto a la enfermedad. Antes de que existiesen tratamientos eficaces contra la enfermedad, los genes hemofílicos, que desde luego se transmiten a la descendencia, se perdían rápidamente, pues pocos niños de familias hemofílicas sobrevivían. Por el contrario, las mutaciones recesivas que afectan a los cromosomas autosómicos (los que no son sexuales), de los cuales hay dos copias en cada célula, pueden propagarse a toda la población, ya que afectan sólo a los contados individuos que hereden los dos genes defectuosos.

En principio cabe identificar la mutación que provoca la hemofilia en un individuo sin más que aislar y secuenciar su gen para el factor VIII. El método, sin embargo, no resulta operativo; llevaría meses secuenciar las 186.000 bases de cada gen. Afortunadamente se dispone de un procedimiento más rápido, si bien sólo aplicable a una pequeña parte de los casos. Se basa en una técnica de hibridación, la denominada transferencia por el método de Southern.

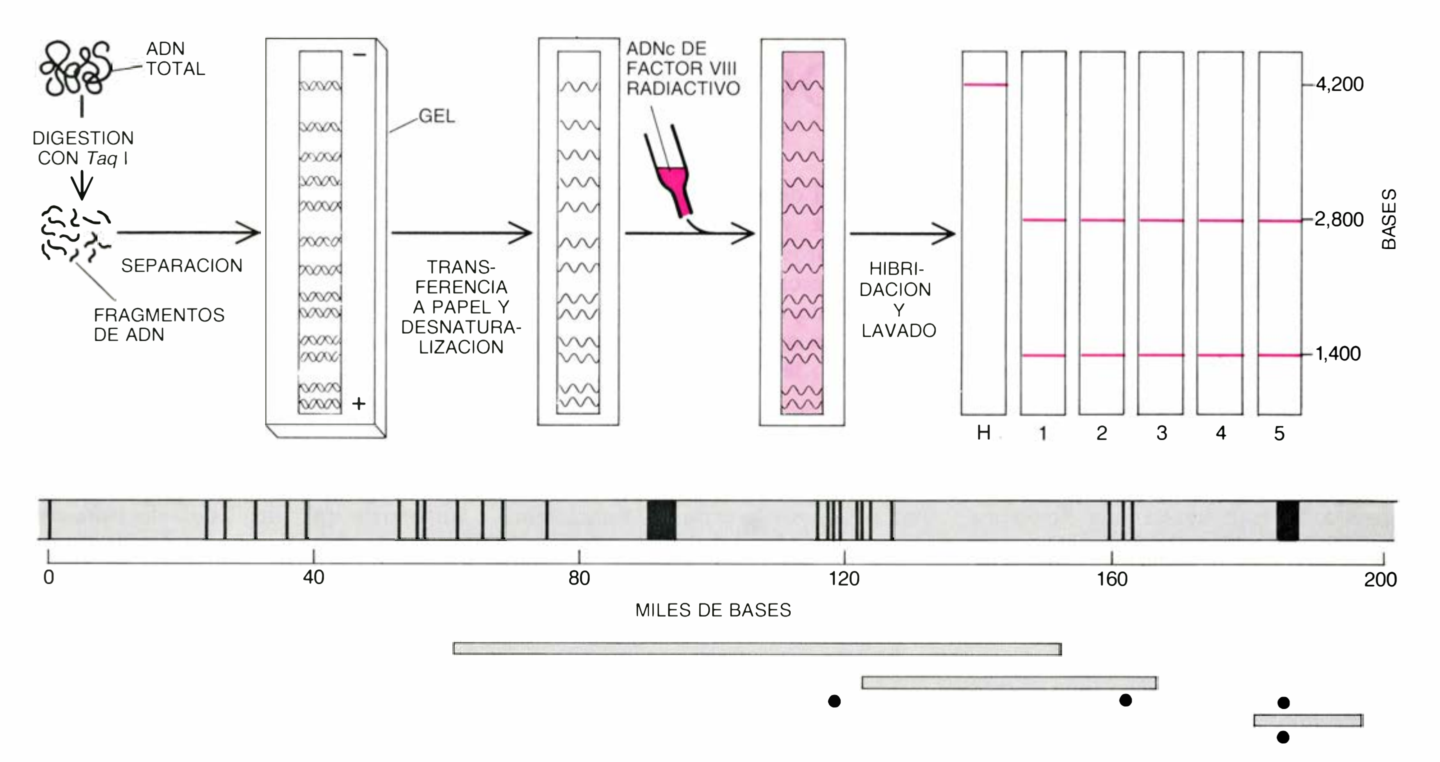
El primer paso contempla la extracción del ADN de las células sanguíneas de un hemofílico. Se corta el ADN en un millón de fragmentos, o más, por medio de una restrictasa, una enzima que corta el ADN cada vez que reconoce una secuencia específica, formada habitualmente por cuatro o seis bases. Se someten los fragmentos a electroforesis, que los separa por tamaños. (En esa técnica, los segmentos menores avanzan más en un gel de agarosa sometido a una corriente eléctrica.) A continuación se separan las cadenas del ADN y se transfieren a un papel de filtro especial, en el que mantienen la misma posición relativa que ocupaban en el gen. Finalmente se sumerge el filtro en una solución que contiene ADNc del factor VIII marcado radiactivamente. La sonda de ADNc hibrida con los fragmentos del gen del factor VIII. Se deducen los tamaños de los fragmentos a partir de la posición que ocupan en el papel, en el que se disponen según un patrón característico.

Para detectar las mutaciones que pudieran afectar al factor VIII, nuestro equipo comparó los patrones de hibridación correspondientes a los ADN de individuos normales y hemofílicos. Con ese método se detectan dos tipos de alteraciones génicas. La más sencilla es alguna deleción importante que pudiera afectar a parte del gen; en ese caso faltarán fragmentos o serán de distinto tamaño, y por tanto variará el patrón de hibridación. También pueden detectarse, a veces, los cambios de una sola de las bases de ADN, siempre que afecten a la secuencia que reconoce la restrictasa utilizada. Esa mutación impide que la enzima corte el gen, y en vez de lo que en el patrón normal de cortes son dos fragmentos, se observará uno solo en el correspondiente al hemofílico. (Y viceversa: puede crear un nuevo sitio de restricción.)

Un ejemplo de mutación de cambio de base bastará para ilustrar el procedimiento. Se corta con la restrictasa *Taq*I el ADN de un individuo con hemofilia, así como el de sus padres y sus tres hermanos. Esa restrictasa reconoce la secuencia *TCGA*. Tras aplicar la técnica de Southern, en los cinco individuos no afectados se observan dos bandas (correspondientes a la hibridación de la sonda con dos fragmentos), una correspondiente a un tamaño de 1400 nucleótidos y la otra de 2800. En el caso del individuo afectado de hemofilia, en vez de dos fragmentos se observa uno solo, de 4200 nucleótidos. Puesto que conocemos la secuencia del gen normal, podemos determinar la posición del sitio *Tag*I alterado. Tras clonar y secuenciar esa parte del gen del individuo hemofílico se advierte que la secuencia que reconoce *Tag*I ha cambiado de *TCGA* a *TTGA*. La mutación impide que *Tag*I corte el punto.




**6. SECUENCIA DE AMINOACIDOS del factor VIII; semeja la de la ceruloplasmina, una proteína de la sangre. Los tres dominios *A* del factor VIII tienen en común una tercera parte de los aminoácidos y son homólogos a los tres dominios de la ceruloplasmina. Las dos proteínas deben haber evolucionado a partir de un antecesor común. Es probable que el factor VIII divergiese de la ceruloplasmina al insertarse en el gen ancestral un gran exón (el número 14), que determina el segmento *B* intermedio. Los exones correspondientes a los segmentos *C* se añadieron a uno de los extremos. Cuando se activa el factor VIII, ciertas proteasas eliminan el segmento *B*, quedando enlazadas las dos subunidades resultantes por un ion de calcio.**




**7. MUTACIONES QUE CAUSAN HEMOFILIA. Situadas en el gen del factor VIII, pueden detectarse mediante la técnica denominada transferencia por el método de Southern (*arriba*) siempre que modifiquen el patrón de fragmentos que se obtiene al cortar el gen con una restrictasa. Se corta el ADN de células sanguíneas en millones de fragmentos, en este caso con la enzima *Taq*I, se separan los fragmentos por tamaños mediante electroforesis, se desenrolla el ADN en sus dos cadenas y se transfiere a un papel de filtro. Se baña el filtro en una solución que contiene ADNc radiactivo del factor VIII, que sólo hibridará con fragmentos del gen del factor VIII. El tamaño de los fragmentos que hayan hibridado se infiere exponiendo una película de rayos X al filtro. En el ejemplo de la figura, una mutación puntual en el gen del factor VIII de un hemofílico (*H*) ha eliminado un sitio de corte de *Taq*I. Los fragmentos de 2800 y 1400 bases, presentes en los filtros correspondientes a los familiares normales (*1-5*), se han sustituido por uno solo, de 4200 bases. Hasta hoy se han localizado siete mutaciones diferentes en genes de factor VIII de individuos hemofílicos (*abajo*). Cuatro son cambios de una sola base (*puntos*) y tres son grandes deleciones (*barras*).**

A la vez, y ello sí es de la mayor importancia, esa mutación ha provocado el cambio del codón para el aminoácido arginina (*CGA*) por otro de "paro" (*TGA*), que provoca el cese prematuro de la síntesis del factor VIII. La proteína así truncada es inactiva o demasiado inestable para sobrevivir en el torrente sanguíneo. Puesto que los padres del individuo hemofílico carecen de la mutación, éste no puede haberla heredado. Debe tratarse de una mutación nueva, ocurrida en el zigoto a partir del cual se desarrollaron todas sus células.

Entre nuestro laboratorio, el Instituto de Genética y la Universidad Johns Hopkins se han examinado hasta la fecha unos 200 genes del factor VIII de individuos hemofílicos. Se han caracterizado siete mutaciones diferentes y ninguna de ellas se ha observado en más de una familia. Cuatro de ellas son cambios de una sola base, de las que tres determinan la síntesis de un factor VIII truncado y provocan una grave hemofilia; la cuarta conlleva el cambio de un aminoácido y una hemofilia moderada. Las otras tres mutaciones son deleciones que afectan a varios miles de nucleótidos del gen; en los tres casos la hemofilia es grave.

En el futuro, probablemente cuenten los investigadores con técnicas más eficaces para localizar mutaciones de cambios de bases. Por el análisis de gran número de mutaciones podrá establecerse una correlación entre los tipos de mutación y la gravedad de la enfermedad. Será de enorme importancia, por ejemplo, dilucidar por qué el diez por ciento de los hemofílicos presentan respuesta inmunitaria frente al factor VIII exógeno. Son esos casos los que más cuesta tratar.

En principio cabría curar a los hemofílicos introduciéndoles genes funcionales del factor VIII en sus células. No obstante, la terapia de genes para cualquier enfermedad se ve muy lejana. Uno de los grandes obstáculos es el problema del control de la productividad de los genes insertados. Un exceso de factor VIII, por ejemplo, puede resultar tan peligroso como su escasez.

Se utiliza ya el gen del factor VIII clonado en el desarrollo de métodos fiables de diagnóstico de hembras portadoras y de detección de la hemofilia en los fetos. En esencia, esos métodos se basan en las pruebas de hibridación, utilizando el gen clonado como sonda para estudiar la herencia de un gen defectuoso. La técnica de diagnóstico prenatal se emplea ya en unos 70 complejos hospitalarios de todo el mundo. Sin ser aún aplicable a todos los casos, resulta más fiable que el método tradicional de medir la concentración de factor VIII en la sangre fetal, y no requiere la práctica de incisión alguna. Además, mientras que la prueba de la sangre fetal no puede realizarse antes de las 20 semanas de embarazo, el diagnóstico por ADN resulta ya viable en la octava semana. Si los padres optan por el aborto, el riesgo que corre la madre es menor.

Para el clínico, la consecuencia más trascendental de la clonación del gen del factor VIII es, obviamente, la perspectiva de un aporte seguro y abundante de factor VIII. La proteína fabricada por ingeniería genética entrará muy pronto en su programa de ensayos clínicos, que durará varios años. Una vez lograda su comercialización, los hemofílicos se habrán liberado de la amenaza que suponen las transfusiones infecciosas. Muchos de quienes habiten países subdesarrollados, donde todavía hoy mueren gran número de jóvenes, podrán recibir por primera vez un tratamiento eficaz.

# Ciencia y sociedad

### *Pauli y la teoría atómica*

A los 21 años, Wolfgang Pauli era considerado ya uno de los talentos más prometedores de su tiempo. Alta estimación que se debía, sobre todo, a Albert Einstein, quien así se pronunció públicamente al leer su exposición de la teoría de la relatividad en la *Enciclopedia de las Ciencias Matemáticas*. El propio Max Born, después de haber estudiado uno de los trabajos de Pauli, escribía: "Especialmente interesante me pareció su conclusión... [según la cual] la aplicación de la teoría del continuo al interior del electrón no tiene sentido tratándose de un hecho por principio inobservable. Justamente esta idea la estoy persiguiendo en vano desde hace algún tiempo; esto significa, a mi parecer, que la solución de todas las dificultades cuánticas ha de buscarse desde un punto de vista básicamente nuevo: los conceptos de espacio y de tiempo, concebidos como un continuo cuadridimensional, no pueden transferirse del mundo macroscópico de nuestra experiencia al mundo de los átomos. Esto requiere, sin duda, como representación adecuada otro tipo de variedad numérica. No sé cómo hacerlo. Aún no soy viejo, pero ya tengo demasiados años y tengo demasiadas preocupaciones como para que se me ocurra. Este es su trabajo, su vocación, de acuerdo con todo lo que he oído de usted." La opinión favorable de Born se robusteció con los años: "Pauli era un genio, sólo comparable con Einstein, y, desde el punto de vista científico, quizá superior a él."

Durante cuatro decenios, de 1918 a 1958, Pauli ejerció una enorme influencia sobre el desarrollo de la física de su tiempo. Aparte de sus publicaciones y de sus clásicos tratados sobre la teoría de la relatividad (1921), la teoría cuántica (1926) y la mecánica ondulatoria (1933), dio grandes impulsos a la física, a través de una extensa correspondencia.

La prioridad de un descubrimiento científico se establece generalmente en razón de su publicación. Sin embargo, para estudiar la génesis de las nuevas ideas el mejor medio son las cartas. Permiten seguir de cerca la evolución de los conceptos en las etapas intermedias y ponen de manifiesto los motivos y las influencias externas que guiaban al descubridor. En el caso de Pauli esta vía es especialmente fecunda, porque muchos de sus resultados no vieron la luz pública: unas veces, porque no deseaba entorpecer el trabajo de otros; otras, porque el resultado no le parecía de suficiente interés.

Veamos como ejemplo su producción literaria durante el año 1926. Publicó sólo dos trabajos. El aparecido en enero, aunque terminado en diciembre del año anterior, abordaba el átomo de hidrógeno de acuerdo con la mecánica matricial descubierta por Heisenberg seis meses antes. Fue el segundo su tratado sobre la teoría cuántica, preparado asimismo en 1925, obra que constituye la última exposición competente de la física atómica antes del advenimiento de la nueva mecánica de Heisenberg. Por tanto, la única publicación correspondiente al año 1926 (aparecida en 1927) es la teoría del paramagnetismo, basada en la aplicación del principio de exclusión a los electrones de una red cristalina, que inauguró la física cuántica del estado sólido.

Entre los trabajos de 1926 no publicados figuran: *1*) El desarrollo de un formalismo para el tratamiento de sistemas no periódicos de acuerdo con la mecánica de Heisenberg. Pauli interrumpió su trabajo al conocer que Dirac seguía una línea similar. *2*) Demostró –independientemente de Schroedinger– la equivalencia de las dos versiones (de Heisenberg y de Schroedinger) de la nueva teoría cuántica. Al comunicar a su amigo Schroedinger los valores de las intensidades de las líneas espectrales del hidrógeno calculados por él, éste le contestó: "Me siento muy ruin al escribirte siempre: lo que me comunicas, ya está en una de mis notas a la prensa." *3*) Según el testimonio de Pascual Jordan, Pauli fue también uno de los primeros que propusieron la interpretación probabilística de la función de ondas de Schroedinger. La idea fue acogida por Max Born en su famoso estudio de las colisiones atómicas de acuerdo con la mecánica ondulatoria. *4*) Pauli intervino en la apasionada disputa que se entabló en otoño de 1926 entre Bohr y Schroedinger en torno a la interpretación correcta de la nueva teoría. Las sugerencias de Pauli ayudaron a Heinsenberg a reconocer la relación de imprecisión que existe entre el impulso $p$ y la posición $q$ de la partícula, en una frase célebre que Pauli comunicó en octubre de 1926 a Heisenberg: "El mundo se puede observar con el ojo $p$ y se puede observar con el ojo $q$. Pero si se intenta abrir a la vez los dos ojos, uno se vuelve loco." *5*) Menos favorablemente acogieron algunos contemporáneos la intervención de Pauli en el debate por el espín electrónico, abierto en octubre de 1925 por los físicos holandeses George Uhlenbeck y Samuel Goudsmit. Pauli se opuso enérgicamente, al enterarse que éstos recurrían a una representación pictórica. Para él, introducir un "giro" del electrón era contrario a la nueva tendencia de la teoría atómica que evitaba todo uso de modelos. Después de arduos debates, Pauli se rindió en marzo de 1926, una vez que su principal objeción (la aparición de un factor 2 inexplicable, en las fórmulas espectrales de los dobletes) había sido aclarada por los cálculos de L. H. Thomas, que por entonces visitaba Copenhague.

La denominación de "electrón magnético", propuesta por Pauli para evitar la imagen de la rotación, fue abandonada en favor de la de espín, introducida por los autores ingleses.

También en este caso las cartas ayudan a comprender mejor los motivos de la oposición de Pauli y su intervención a la hora de aclarar e integrar el nuevo concepto de espín en la teoría atómica. Las cartas no sólo son un instrumento de reconstrucción del desarrollo histórico, sino también de corrección de tradiciones erróneas implantadas durante las diversas etapas evolutivas de la ciencia.

Pauli vivió su juventud en Viena en un ambiente de gran turbulencia política y cultural. Durante los últimos años escolares había adquirido el dominio de las matemáticas superiores, dirigido por Ernst Mach, cabeza del positivismo moderno, amigo del padre de Wolfgang y padrino del hijo. Esta relación iba a reflejarse en la actitud de Pauli. Mach seleccionaba la literatura científica adecuada para el joven precoz. Compartió el aula con otro alumno sobresaliente, Richard Kuhn, quien posteriormente obtuvo el Nobel de química.

Aunque la Universidad de Viena disponía de un excelente claustro docente, cuando terminó sus estudios medios en 1918 decidió acudir a la de Munich. Aquí Arnold Sommerfeld había creado un centro de física teórica de fama internacional. Bajo su dirección, Munich





*1. Wolfgang Pauli y Albert Einstein en Leiden en el otoño de 1926. (La foto fue tomada por Paul Ehrenfest)*



*2. Pauli en sus tiempos de estudiante en el "Doeblinger Gymnasium" en Viena*

se había convertido, con Copenhague, en la meca de la física atómica.

Partiendo de los trabajos fundamentales de Niels Bohr sobre la cuantificación del átomo de hidrógeno, Sommerfeld había generalizado el procedimiento para sistemas atómicos más complicados, como son los sistemas múltiplemente periódicos. (En tales sistemas –después de haber expresado adecuadamente sus coordenadas correspondientes a sus grados de libertad– el movimiento de las partículas constituyentes puede considerarse una superposición de movimientos armónicos simples.) Esto era la base de la así llamada antigua teoría cuántica de Bohr-Sommerfeld, que hasta su sustitución por la mecánica cuántica en 1925 dominaba la física atómica.

Preocupaba entonces la búsqueda de modelos atómicos apropiados, susceptibles de abordarse como sistemas múltiplemente periódicos. Un primer triunfo de esta teoría fue la determinación de la estructura fina de las líneas espectrales, realizada en 1915 por Sommerfeld, como efecto de la variabilidad relativista de la masa electrónica.

Durante sus tres años de estudio en Munich, de 1918 a 1921, Pauli hizo importantes contribuciones. Introdujo como unidad cuántica de momento magnético el magnetón de Bohr. Presentó la innovación en septiembre de 1920 en Bad Nauheim, durante la primera asamblea de científicos alemanes celebrada después de la guerra. Allí tuvo lugar también la memorable disputa entre Einstein y Philipp Lenard sobre la teoría de la relatividad.

Sommerfeld propuso a Pauli trabajar, como tesis de doctorado, en la cuantificación de la molécula-ion $H_2^+$. Era uno de los cálculos más ambiciosos dentro del esquema de la antigua teoría cuántica, a pesar de ser el sistema bicéntrico más simple. Su tratamiento requería técnicas avanzadas de mecánica analítica, desarrolladas en el siglo anterior por Carl Gustav Jacobi y Henri Poincaré, y una ampliación del principio de correspondencia de Bohr.

Pauli sufrió su primera desilusión al comprobar que los cálculos no se ajustaban a los datos empíricos. Pero el tratamiento del problema había sido riguroso y, por tanto, este resultado constituyó un primer caso donde la teoría cuántica fallaba. Su amigo Werner Heisenberg, al enterarse del fracaso, añadió simbólicamente una cruz (en vez del signo positivo) al símbolo $H^+$.

Con ello se inició la crisis de la antigua teoría cuántica, a pesar de que su consumación tardó más de un año. La superación llegó a fines de 1924 y en el verano de 1925, con las contribuciones de dos alumnos de Sommerfeld: el principio de exclusión de Pauli y la mecánica matricial de Heisenberg.

En 1922, cuando Pauli preparaba la publicación de su tesis, aún se confiaba en salvar la teoría. Se creía que la aplicación del cálculo de perturbaciones remediaría los defectos. Max Born, poco después de ocupar la dirección del nuevo instituto de física teórica creado en Gotinga, nombró a Pauli su asistente para el semestre de invierno 1921-22. Con su ayuda Born adaptaba

los métodos de perturbación de los astrónomos a las exigencias de la teoría atómica. El objetivo era tratar ahora sistemas atómicos con un número mayor de electrones.

En vista de la gran complejidad, Sommerfeld y sus discípulos habían adoptado en sus modelos atómicos diversas configuraciones electrónicas, guiándose por su intuición y por su creencia pitagórica en las armonías predilectas de la naturaleza. La nueva tendencia aspiraba, en cambio, a remplazar estas conjeturas arbitrarias por sistemas cuyas configuraciones electrónicas se obtenían sistemáticamente por aplicación de las leyes de la mecánica suponiendo interacciones electrodinámicas entre las partículas constituyentes del átomo.

Diversos autores, que aplicaron estos nuevos métodos al átomo de helio con dos electrones, llegaron a la misma conclusión fatal que Pauli ya había experimentado anteriormente con el $H_2^+$. Con ello, durante el año 1923, la convicción de la crisis se hizo general. Los más adelantados y revolucionarios eran los jóvenes, que creían que la reforma sólo podría conseguirse abandonando el esquema de la física clásica, basado en modelos atómicos y la descripción espacio-temporal de los fenómenos.

En Gotinga existía el convencimiento de que la búsqueda sistemática del origen de las discrepancias en la física atómica conduciría a la teoría definitiva. Born hablaba de una "discretización" de las leyes atómicas, que hasta entonces se habían expresado mediante ecuaciones diferenciales en el contexto de una teoría del continuo. Bohr, en cambio, empleaba simultáneamente los datos que le proporcionaban los espectros de los elementos en el sistema periódico de los elementos y los principios heurísticos que le habían prestado grandes servicios anteriormente.

De especial utilidad se mostró su principio de correspondencia, que permitía establecer una relación estrecha entre los fenómenos clásicos y cuánticos. Así, Bohr obtuvo las reglas de selección, las de polarización y fórmulas para las intensidades de las líneas espectrales emitidas por los átomos, considerando las oscilaciones armónicas correspondientes a los estados estacionarios del sistema involucrados en las transiciones respectivas. "Yo creo realmente –escribía Paul Ehrenfest a su amigo Einstein en diciembre de 1921 con admiración del trabajo de Bohr– que él domina la estructura de todos los átomos, con sólo muy pequeñas incertidumbres en algunas pocas sustancias; y esto casi únicamente por medio de consideraciones cualitativas (¡sutiles!) y cálculos estimativos. (Sólo He requiere cálculos más profundos.)" Bohr proponía luego estudiar la ordenación de los elementos en el sistema periódico y su comportamiento en campos eléctricos y magnéticos.

Mientras tanto, se acercaba el fin de la estancia de Pauli en Gotinga. Acostumbrado a la vida de la gran ciudad, no se encontraba allí a gusto. Además, Max Born era hombre demasiado formal y meticuloso como para comprender la indolencia y la falta de respeto de su colaborador, que solía levantarse al mediodía y llegaba con retraso a las reuniones convocadas por la mañana.

En abril de 1922 aceptaba el puesto de colaborador científico del nuevo director del Instituto de Física Teórica de la Universidad de Hamburgo, Wilhelm Lenz. Pauli se incorporó allí en un grupo de físicos y matemáticos jóvenes, dispuesto a una vida más libre y desligado de los hábitos convencionales. Todavía comentan los físicos muchas anécdotas relacionadas con efectos extraños y físicamente inexplicables, los "efectos Pauli", que acontecían por la sola presencia de éste y que tuvieron su origen en ese ambiente.

Pronto se les unió Otto Stern, nombrado en 1923 director del Instituto de Química Física. Con él y con Walther Gerlach, Pauli ya había establecido contactos anteriormente, cuando éstos efectuaron su experimento que puso de



*3. Pauli con su profesor Arnold Sommerfeld en el congreso sobre metales desarrollado en octubre de 1934 en Ginebra*

manifiesto la cuantificación direccional de las órbitas electrónicas en campos magnéticos. Hamburgo se convirtió así, durante algunos años, en importante centro de la nueva física.

La ciudad quedaba más cerca de Copenhague. Su primer encuentro con Niels Bohr tuvo lugar durante las célebres jornadas bohrianas en Gotinga, en junio de 1922. Con sus preguntas y observaciones hechas durante excursiones realizadas después de las conferencias, Werner Heisenberg y Wolfgang Pauli impresionaron a Bohr. Pauli fue invitado a Copenhague el otoño de 1922.

Allí tuvo Pauli la oportunidad de conocer más de cerca las ideas de Bohr. También el danés estaba muy contento con su nueva adquisición, según decía, en una carta fechada en febrero de 1923 y dirigida a Ehrenfest: "Me alegro de que usted esta vez esté menos descontento que con mi último trabajo; quizás esto se deba también a la presencia de Pauli, que me ha ayudado bastante en la redacción. Todos nosotros le queremos mucho y recientemente he escrito con él un pequeño trabajo sobre los temas generales que se presentan en la interpretación cuántica del efecto Zeeman. Estoy muy impaciente por oír su opinión y de saber si usted lo encuentra demasiado conservador o demasiado revolucionario."

Ya que en el caso del efecto Zeeman anómalo la separación de las líneas espectrales en un campo magnético no se ajustaba a las predicciones de la teoría, los físicos suponían, con razón, que la elucidación de este fenómeno podría señalar el camino hacia la teoría atómica correcta. Pero el esfuerzo conjunto de Bohr y Pauli por solucionar el problema fue en vano. Hoy sabemos el motivo.

Según la concepción de aquel tiempo, las capas electrónicas completas llevaban asociadas un momento magnético. Los multipletes de las líneas espectrales se creían originados por las diferentes orientaciones del momento magnético asociado a la órbita del electrón de valencia con respecto al momento magnético del residuo atómico ("giróscopo interior"). Pauli demostró que el verdadero origen de la duplicidad de los espectros alcalinos se debía al electrón de valencia.

La imposibilidad de resolver estos problemas aumentaba la desconfianza en los modelos atómicos. Pauli decidió formular en adelante todos los resultados obtenidos sin recurrir a modelos. "Con todos los modelos uno tiene la fuerte impresión –escribía un diciembre de 1924 a su profesor Sommerfeld– que estamos hablando un idioma que no se adapta lo suficiente a la sencillez y la armonía del mundo cuántico."

Primer fruto de su renuncia, logró establecer una conexión entre las líneas Zeeman que aparecen en campos magnéticos fuertes y débiles.

Al regresar a Hamburgo en el otoño de 1923, dejó abandonada por algún tiempo la física atómica. "Corría agitado y con una cara huraña por las calles de Copenhague –recordaba posteriormente Pauli– pensando en el efecto Zeeman anómalo." En sus momentos de desilusión sobre el estado de la física atómica, Pauli solía entretenerse con un problema de física clásica: la conducción térmica. "Más tarde, quizás muy pronto, volveré a preocuparme de [la física atómica] con nuevos esfuerzos", escribía en febrero de 1924 a Bohr.

La ocasión se presentó en su lección inaugural de Hamburgo, al hacerse cargo de una plaza de "docente privado". Como tema había elegido la exposición de las ideas de Bohr acerca de la constitución del sistema periódico de los elementos. A partir de trabajos anteriores de J. J. Thomson, W. Kossel y A. Sommerfeld, Bohr había logrado ordenar todos los electrones de los diversos elementos en grupos, usando los números cuánticos orbitales $n$ y $k$, que determinan los dos ejes (mayor y menor, respectivamente) de la órbita elíptica del electrón correspondiente. Mas para la longitud de los períodos del sistema periódico, la teoría de Bohr no daba ninguna explicación satisfactoria. Esta deficiencia, muy criticada por Pauli, fue el punto de partida de su intervención. Por otra parte, "Sommerfeld en su libro *Constitución atómica y líneas espectrales* había llamado 'cabalística' la famosa fórmula $2p^2$, siendo $p$ un número entero. Esto me había impresionado profundamente cuando yo era estudiante", recordaba Pauli.

Pauli se opuso a los intentos de Bohr por explicar esta fórmula a través del principio de correspondencia. Acudiendo a este principio Bohr creía que se podría presentar algún día la teoría cuántica como generalización racional de la teoría clásica. Puesto que, mediante ella, se podían comprender las reglas de selección espectroscópicas, Bohr también esperaba justificar la conclusión de los períodos en el sistema periódico.

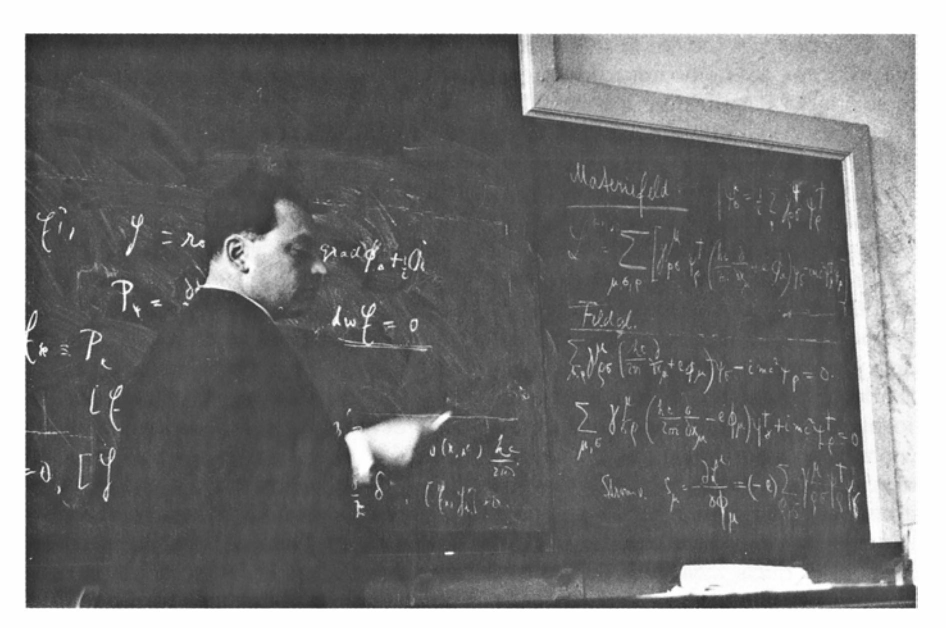
La raíz de todas estas dificultades era el uso de concepciones atómicas erróneas y de la aplicación de una teoría falta de coherencia. La solución llegó a través de tres etapas: localización del origen de la anomalía magnetomecánica, hallazgo de la ley de distribución de los electrones entre los diferentes estados cuánticos y formulación de una mecánica cuántica correcta.

El modelo del "giróscopo interior" lo había usado en 1921 Werner Heisenberg para explicar la duplicidad de las líneas espectrales en los álcalis. De acuerdo con la cuantificación espacial, se admitían sólo dos orientaciones de los momentos angulares que intervenían.

Al clasificar los términos espectrales mediante los números cuánticos introducidos por Bohr y Sommerfeld, Alfred Landé, de la Universidad de Tubinga, se vio obligado a introducir en sus expresiones números cuánticos semi-enteros. Eso se oponía al espíritu de la teoría. "Los físicos atómicos alemanes –resumía Pauli en febrero de 1924– se dividen ahora en dos clases. Unos calculan primero un problema usando números cuánticos semi-enteros, o si el resultado no se ajusta a la realidad, calculan finalmente con número cuánticos enteros... Este tipo de física teórica me desagrada. Prefiero dedicarme a mi problema de la conducción térmica en el cuerpo sólido... Soy mucho más radical que los físicos atómicos 'semi-enteros'. Pues no creo que las desviaciones de la realidad, señaladas por los resultados al aplicar la teoría de los sistemas periódicos a sistemas atómicos con varios electrones, se eliminen con sólo introducir números semi-enteros en vez de enteros en las fórmulas finales." Otro problema de la teoría espectral aparentemente conectado era la violación del teorema de Larmor. Siendo $J$ y $M$ el momento mecánico y magnético correspondiente a una órbita electrónica, este teorema establecía entre ambos la relación giromagnética $\gamma = M/J = e/2mc$, donde $e$ y $m$ designan las constantes de carga y masa del electrón y $c$ la velocidad de la luz. Al determinar los momentos magnéticos asociados a las diferentes órbitas, que dan origen a los términos espectrales correspondientes al efecto Zeeman anómalo, Landé encontró para el valor de $\gamma$ ciertas desviaciones características: para el "giróscopo interior" el valor de $\gamma$ era el doble de lo que le correspondía según el teorema.

Esta misma anomalía había sido descubierta en otro contexto por el físico suizo Emil Beck, al repetir el experimento de Einstein-de Hass para comprobar la existencia de las corrientes moleculares amperianas.

Ante estas dificultades, Pauli pro-



*4. Pauli en abril de 1929 exponiendo durante la conferencia de física en Copenhague su trabajo con Heisenberg "sobre la electrodinámica cuántica de los campos de onda". [Las ecuaciones del pizarrón son las mismas que aparecen en la publicación en la* Zeitschrift für Physik, 56*, 1-61 (1929)]*

puso la siguiente reforma, según comunicaba con cierto orgullo a Bohr en diciembre de 1924: "Al redactar [en un artículo sobre la teoría cuántica] la teoría del efecto Zeeman y el teorema de Larmor, se me ocurrió que, de acuerdo con la teoría cuántica normal, la dependencia relativística de la velocidad por parte de la masa del electrón debería ya mostrar desviaciones de este teorema."

Según la teoría convencional, la anomalía magnética se suponía localizada en la capa electrónica interior más cercana al núcleo atómico; allí, en los átomos de número atómico elevado, los electrones adquieren velocidades relativísticas. Por tanto, también la relación giromagnética, en que interviene la masa del electrón, debería revelar esta dependencia. "En cambio –concluía Pauli– las observaciones del efecto Zeeman son realmente incompatibles con el factor de corrección... Escapatorias hay varias, pero yo no creo en ellas. Creo que se exageró la realidad de los impulsos interiores en las capas electrónicas (inertes). Además, me parece insostenible suponer que la anomalía del efecto Zeeman se deba al magnetismo doble del residuo atómico. Esta hipótesis (que en el fondo fue propuesta por mí y que se había tomado demasiado en serio por los otros), se opone vivamente a mi instinto físico... En cambio he intentado llevar lo más lejos posible la siguiente hipótesis de trabajo. La estructura de los dobletes alcalinos es, en esencia, una mera propiedad del electrón de valencia. Esta misma duplicidad, incomprensible desde el punto de vista clásico y causa de la estructura de los dobletes, origina también la anomalía de los efectos Zeeman."

Con esta aclaración, el estudio de los efectos Zeeman anómalos quedaba reducido al estudio de los electrones de valencia. Pauli decía que había eliminado así el "giróscopo en el estómago de cada átomo, como la manzana del cuento de Blancanieves".

El escepticismo mantenido hacia los modelos atómicos por los físicos cuánticos de su generación le libraba de interpretar esta nueva y curiosa duplicidad del electrón como un nuevo grado de libertad de movimiento rotativo. Su oposición a la hipótesis de Uhlenbeck y Goudsmit, sin embargo, obtuvo cierta justificación, al revelarse posteriormente en la teoría de Dirac que el espín era una verdadera propiedad relativista. El carácter esencialmente no clásico del espín se manifestó también en la imposibilidad de su medición en electrones libres mediante un experimento clásico.

Al ocuparse de la formación de los períodos en el sistema periódico de los elementos en su artículo enciclopédico mencionado, Pauli consultó la nueva edición de 1924 de la obra de Sommerfeld. Esta aludía, en el prefacio, a una nueva clasificación de los subniveles atómicos, propuesta por el físico inglés Edmund Stoner. "Apenas había leído su prefacio –le comentaba Pauli inmediatamente– corrí a la biblioteca para estudiar el trabajo de Stoner. Estoy convencido de que la modificación del esquema de Bohr propuesta por él es la correcta."

Para Stoner, además de completarse una capa de $n$ con $2n^2$ electrones, cada subgrupo, caracterizado por $n$, $k$ según la clasificación de Bohr, sólo puede tener como máximo $N(k) = 2\ (2k-1)$ electrones. Stoner descubrió que esta ocupación máxima de los subgrupos era igual al número de niveles energéticos del electrón correspondiente al mismo $k$ en un campo magnético externo.

Pauli advirtió que estás proposiciones podían ser generalizadas, "de modo que no sólo se pueden fijar los números electrónicos en las capas completas, sino también *el posible número de realizaciones* (valores de $j$) *de las capas incompletas* con un número determinado de electrones". Si esta generalización se verificaba también en el caso de espectros más complicados, explicaba a Sommerfeld, "entonces usted tenía toda la razón al confiar 'más en el poder mágico de los cuantos que en las consideraciones de correspondencia y de estabilidad'".

Al volver de sus vacaciones navideñas, que había pasado con sus padres en Viena, Pauli hizo una breve visita a la Universidad de Tubinga en enero de 1925. Heisenberg le había informado que el famoso espectroscopista Ernst Back "escondía" allá sus nuevas mediciones de las líneas Zeeman del plomo. Tras comprobar que todas las líneas estaban de acuerdo con su principio, Pauli regresó a Hamburgo y procedió inmediatamente a la publicación del trabajo. Bohr había aprobado también el trabajo; decía que con él Pauli había logrado "la deseada movilización de nuestras ideas; yo tengo la impresión de que hemos llegado al punto crítico decisivo".

En su enunciado definitivo, el principio está expresado en términos de los cuatro números cuánticos, que Pauli había utilizado anteriormente para caracterizar un estado cuántico en campos magnéticos débiles y fuertes. En una formulación que ya hace uso de la nomenclatura del espín $s$ y del momento angular $l$ (y donde $m_s$ y $m_l$ son las correspondientes proyecciones en la dirección del campo). Este principio dice: "Supóngase que actúe un campo magnético tan intenso que se establezca una cuantificación de acuerdo con los números cuánticos $n^{(p)}$, $l^{(p)}$, $m_l^{(p)}$, $m_s^{(p)}$ de cada electrón $p$. *Nunca puede ocurrir entonces que todos los valores de los cuatro números cuánticos sean los mismos para dos electrones en un mismo átomo: o, lo que es lo mismo,*

*un estado del átomo caracterizado por los valores de $n^{(p)}$, $l^{(p)}$, $m_l^{(p)}$, $m_s^{(p)}$ está ya 'ocupado' por un solo electrón*. Esta regla también puede denominarse la 'prohibición de electrones equivalentes' o, brevemente, 'regla de exclusión'."

Los contemporáneos preferían hablar, sin embargo, de la regla de prohibición de Pauli, que según él no era susceptible de una explicación mediante conceptos clásicos. "Yo sospecho –decía en diciembre a Bohr– que la futura justificación de la regla hay que buscarla en otra dirección. Para personas débiles, que necesitan la imagen de órbitas electrónicas y de modelos mecánicos, unívocamente definidas, la regla pudiera justificarse así: 'si varios electrones situados en campos intensos poseen los mismos números cuánticos, tendrían la misma órbita y chocarían; y estos casos hay que excluirlos'... Sin embargo, resultaría mucho más satisfactorio que se lo pudiese comprender directamente en virtud de una mecánica cuántica más general, en la cual estos estados no sean admitidos como estacionarios. Probablemente, al tener esta comprensión más directa para el átomo de hidrógeno, la justificación de la regla de exclusión en los grupos dedúzcase de forma natural. Por otro lado, no creo que esta comprensión pueda obtenerse mediante el principio de correspondencia. Espero una aclaración mediante un análisis físico de los conceptos de movimiento y de fuerza en el sentido de la teoría cuántica, guiada por los hechos empíricos. ¿Estará el tiempo ya maduro para ello?"

Esta declaración, hecha un año antes del descubrimiento de la verdadera mecánica cuántica, demuestra la estrecha relación que existía entre las aspiraciones de Pauli y de Heisenberg con respecto a la reforma de los fundamentos teóricos de la física atómica.

A pesar de que Pauli reconocía que, desde el punto de vista lógico, era legítimo preguntar por una "razón natural" para la aparición de los períodos, su intuición física se oponía a esta posibilidad. Bohr, en cambio, insistía en su posición al hablar de un sistema "natural" cuando se refería al sistema periódico de los elementos.

La nueva mecánica tampoco dio la respuesta que Pauli esperaba. Así "cayó una cierta sombra de incompletitud sobre la luz brillante que emanaba de los éxitos de la nueva mecánica cuántica", confesaba Pauli, al recibir el premio Nobel del año 1945 por el descubrimiento de su principio.

Cuando en octubre de 1931 se le otorgó la medalla Lorentz por el descubrimiento de su principio, su amigo, Paul Ehrenfest, destacó la importancia que había adquirido en pocos años para toda la física: "Los efectos de este principio tan esotérico pueden percibirse incluso en medio de nuestras experiencias diarias. Es delicioso ilustrarlo con un ejemplo sencillo. Tomemos un trozo de metal, o una piedra. Por poco que reflexionemos, nos sorprenderá que este cuanto de materia no ocupe un espacio mucho más reducido. Las moléculas naturalmente se apiñan densamente, unas sobre otras. Lo mismo sucede con los átomos en la molécula. –Bien– ¡Pero, ¿a qué se debe, que los átomos sean tan gordos?!... La respuesta es: sólo a la prohibición de Pauli: ¡dos electrones nunca deben hallarse en el mismo estado cuántico! Por este motivo, los átomos son tan gordos, por este motivo la piedra, el trozo de metal, etc., son tan voluminosos... Pero no sólo la constitución de los átomos, también el sistema periódico de los elementos y la estructura de los espectros obedecen a la prohibición de Pauli. Coqueta y fantasmagórica, la prohibición de Pauli se esconde detrás de muchos otros fenómenos de la física y de la química. Por ejemplo, nos ofrece dos modificaciones de la molécula de hidrógeno (orto y parahidrógeno); sin exagerar puede decirse que el uso de las barras de valencia en la química es un resultado de la prohibición de Pauli. El principio regula también las propiedades magnéticas de la materia, la conducción eléctrica y el calor específico de los metales".

Una vez desenredados los problemas que provenían del uso de modelos atómicos inadecuados, el camino quedaba expedito para llegar al formalismo buscado. Max Born, cabeza del grupo de Gotinga, se había empeñado, desde 1924, en dar el "primer paso hacia la mecánica cuántica", proponiendo "leyes rigurosas para números cuánticos arbitrarios", al formular ecuaciones de diferencial más en armonía con el carácter discontinuo de los fenómenos cuánticos.

Bohr buscaba, por otro lado, la reforma basándose en la nueva teoría de la radiación, que había planteado en enero de 1924 junto con su colaborador Hans Kramers y su huésped John C. Slater. Según esta nueva concepción, cada átomo engendraba un campo "virtual de radiación", cuyas frecuencias correspondían a las posibles transiciones cuánticas. Este campo se decía que era virtual, porque su presencia sólo servía para determinar de antemano las probabilidades de transición. Así pretendía conseguir la compatibilidad entre las ondas electromagnéticas continuas emitidas o absorbidas y los correspondientes saltos cuánticos discretos del átomo. Una consecuencia era la no-conservación de la energía y del impulso en cada proceso.

Pauli rechazaba esta teoría desde su comienzo y acusaba a Bohr de difundir doctrinas heréticas entre los físicos. Una vez demostrado el cumplimiento de las leyes de conservación en cada acto elemental, mediante experimentos de coincidencia de Walther Bothe y Hans Geiger, la teoría fue abandonada también por Bohr en la primavera de 1925. La idea de los campos virtuales renació posteriormente en otro contexto, con las ondas probabilísticas de Schroedinger.

Heisenberg, que estuvo en Copenhague y volvía a Gotinga para el semestre de verano de 1925, sintetizó estas tendencias de ambas escuelas en su célebre "reintrepretación de las relaciones cinemáticas y mecánicas", que inauguró el comienzo de la nueva mecánica cuántica.

Sin duda, el esfuerzo de Heisenberg, empeñado en usar sólo magnitudes observables, deriva también de la influencia de Pauli, quien había criticado durante los últimos años el recurso a conceptos clásicos en la teoría atómica, a pesar de que las órbitas electrónicas nunca se habían observado de forma directa.

"La meta aún no alcanzada –había escrito en diciembre del año anterior– ha de ser la de deducir los valores de la energía y el impulso y de todas las otras propiedades de los estados estacionarios físicamente reales y observables, a partir de los números enteros y de las leyes de la teoría cuántica. Pero no debemos forzar los átomos con las cadenas de nuestros prejuicios (a los que pertenece la supuesta existencia de las órbitas electrónicas en el sentido ordinario de la cinemática); debemos, por el contrario, adaptar nuestros conceptos a la experiencia."

Durante su breve visita a Copenhague en marzo de 1925, Pauli seguía atacando la teoría de Bohr-Kramer-Slater. Bajo su "influencia catalítica" y la prueba experimental del ensayo mencionado de Bothe y Geiger, Bohr finalmente cedió: "Ahora no nos queda otra cosa que eliminar todos los rastros de nuestro intento revolucionario", escribía Bohr a su amigo James Franck en Gotinga.

Con la conversión de Bohr, quedó eliminado el último obstáculo que frenaba el desarrollo de métodos adecua-

dos para determinar las intensidades de las líneas espectrales con ayuda de una forma más restringida del principio de correspondencia.

Heisenberg se había dedicado a este problema, intentando solucionarlo para el átomo de hidrógeno. Llegó finalmente a una solución para el caso del oscilador anarmónico, según le comunicara a Pauli en una carta del 24 de junio de 1925.

En esta fase del desarrollo de la mecánica cuántica, la intervención de Pauli fue sólo indirecta. Estaba absorbido en un artículo sobre la teoría cuántica para el *Handbuch der Physik*. Sin embargo, Pauli y Heinsenberg mantenían un fluido intercambio epistolar, donde discutían los progresos que el segundo había obtenido. De esta correspondencia con Heisenberg desgraciadamente se han perdido la mayoría de las respuestas de Pauli. Sólo se puede reconstruir su contenido a través del epistolario de Heisenberg.

Cuando Heisenberg, durante su estancia en Heligoland en junio de 1925, encontró la formulación decisiva de su nueva mecánica, hizo una visita a Pauli en su instituto hamburgués antes de volver a Gotinga. Pauli le recibió con gran entusiasmo animándole a continuar en la misma línea.

"Lo que más me alegró en las consideraciones de Heisenberg es el método de su procedimiento y la tendencia en que se basa", informaba Pauli a su amigo Kramers. "En conjunto creo que, en cuanto a mis concepciones científicas, me encuentro ahora muy cercano a Heisenberg, y en casi todo tenemos las mismas opiniones, en la medida en que esto es posible para dos hombres que piensan de forma independiente."

Así Pauli fue también el primero que utilizó el nuevo formalismo desarrollado por Heisenberg para aplicarlo al problema del átomo de hidrógeno, después de haber concluido el mencionado artículo. Con la publicación de sus resultados en enero de 1926, la mecánica de Heisenberg había pasado el test más crítico, necesario para su aceptación por una generación de físicos no familiarizados con los métodos matriciales empleados en ella. Característica es la reacción de Einstein, al hablar de un "alfabeto brujo", que se protege por su complejidad contra toda refutación.

La aparición de la mecánica ondulatoria de Schroedinger pocos meses después y la comprobación de su equivalencia matemática con el formalismo matricial facilitaron la rápida aceptación de la doctrina. (Karl v. Meyenn.)

# El supercolisionador superconductor

*Hacia 1995 podría estar construido un acelerador 20 veces más potente que el mayor de los hoy existentes. Analizaría la materia con un detalle sin precedentes y recrearía las condiciones que acompañaron el instante inicial del tiempo*

J. David Jackson, Maury Tigner y Stanley Wojcicki

Estamos en el año 1995. Una zona de pastoreo, una explotación agraria o una pradera apenas si dejan entrever que, bajo su piel, corre un túnel en forma de anillo de 85 kilómetros de longitud y de luz suficiente para andar por su interior. En el interior del túnel, un pequeño tranvía facilita las tareas de mantenimiento de dos tuberías criogénicas de algo más de medio metro de diámetro cada una. Dentro de cada conducción hay otro tubo de vacío, mucho menor, por donde discurre un haz de protones mantenidos en su trayectoria mediante potentes imanes superconductores que rodean el tubo. En cada vuelta al anillo, la energía de los protones de los dos tubos del haz se incrementa mediante un pulso de radiondas. En 15 minutos, los protones se aceleran, a lo largo del anillo y en direcciones opuestas, más de tres millones de veces.

De repente se abren unas puertas electromagnéticas y las trayectorias de los haces se ven obligadas a cruzarse. Los protones chocan a pares y parte de la energía de la colisión puede transferirse a una región de diámetro 100.000 veces menor que el diámetro de un protón, a un ritmo que excede, en mucho, a la potencia instantánea de salida de todas las centrales energéticas de la Tierra. Allí, durante un intervalo brevísimo –que es al segundo lo que un segundo es a 100.000 veces la edad del universo–, tendremos una visión instantánea del universo en el momento de su creación. La energía que se concentrará en aquella región y en aquel instante sólo se encuentra hoy entre los rayos cósmicos menos habituales. Tal concentración de energía constituía, sin embargo, el estado del universo transcurridos $10^{-16}$ segundos después de la gran explosión. Nuevas partículas elementales que podrían materializarse a partir de esa energía permitirían elaborar una explicación del origen de la masa.

Vale la pena ponderar la posibilidad de que semejante visión se halle al alcance de la tecnología del siglo XX. El túnel, la conducción, los sistemas operacionales necesarios y un complemento inicial de detectores de partículas y de ordenadores pueden construirse perfectamente con los medios disponibles –si bien en una escala que hasta ahora no se había intentado nunca–; el coste de la inversión se cifra en unos 4000 millones de dólares, en dólares constantes de 1986. El diseño fundamental de los instrumentos está ya probándose a una escala de energía cercana a la vigésima parte: el modelo a escala es el Tevatrón, acelerador de partículas del Laboratorio Fermi (Fermilab). La versión de tamaño real posibilitaría el estudio de procesos energéticos que no son accesibles a ninguno de los aceleradores hoy en funcionamiento, ni proyectados en parte alguna del mundo.

La máquina propuesta incorpora la experiencia acumulada durante más de 50 años de diseño y construcción de aceleradores de partículas. Desde 1982, además, la comunidad de físicos de altas energías ha dedicado considerables esfuerzos al diseño de una máquina asequible, en términos económicos, que contara, al propio tiempo, con las mayores posibilidades de resolver las cuestiones teóricas planteadas sobre los constituyentes últimos de la materia. Por responder al diseño de un acelerador de haces colisionantes y considerando el papel que los imanes superconductores desempeñan en la minimización del consumo de energía, se ha denominado al proyecto Supercolisionador Superconductor, o SSC.

Los aceleradores de partículas pueden desencadenar las reacciones más energéticas que quepa estudiar bajo condiciones controladas. Debido a la equivalencia entre masa y energía, la máxima energía de reacción fija la máxima masa observable de las entidades materiales básicas que pueden crearse en el laboratorio. Por tanto, el diseño de un acelerador de partículas determina los límites del conocimiento directo y experimental acerca de la estructura fundamental de la materia. La energía y la masa máximas observables en un acelerador de partículas dependen de la energía de los haces de partículas aceleradas, de la manera como la energía de los haces se libera y de la intensidad de los haces.

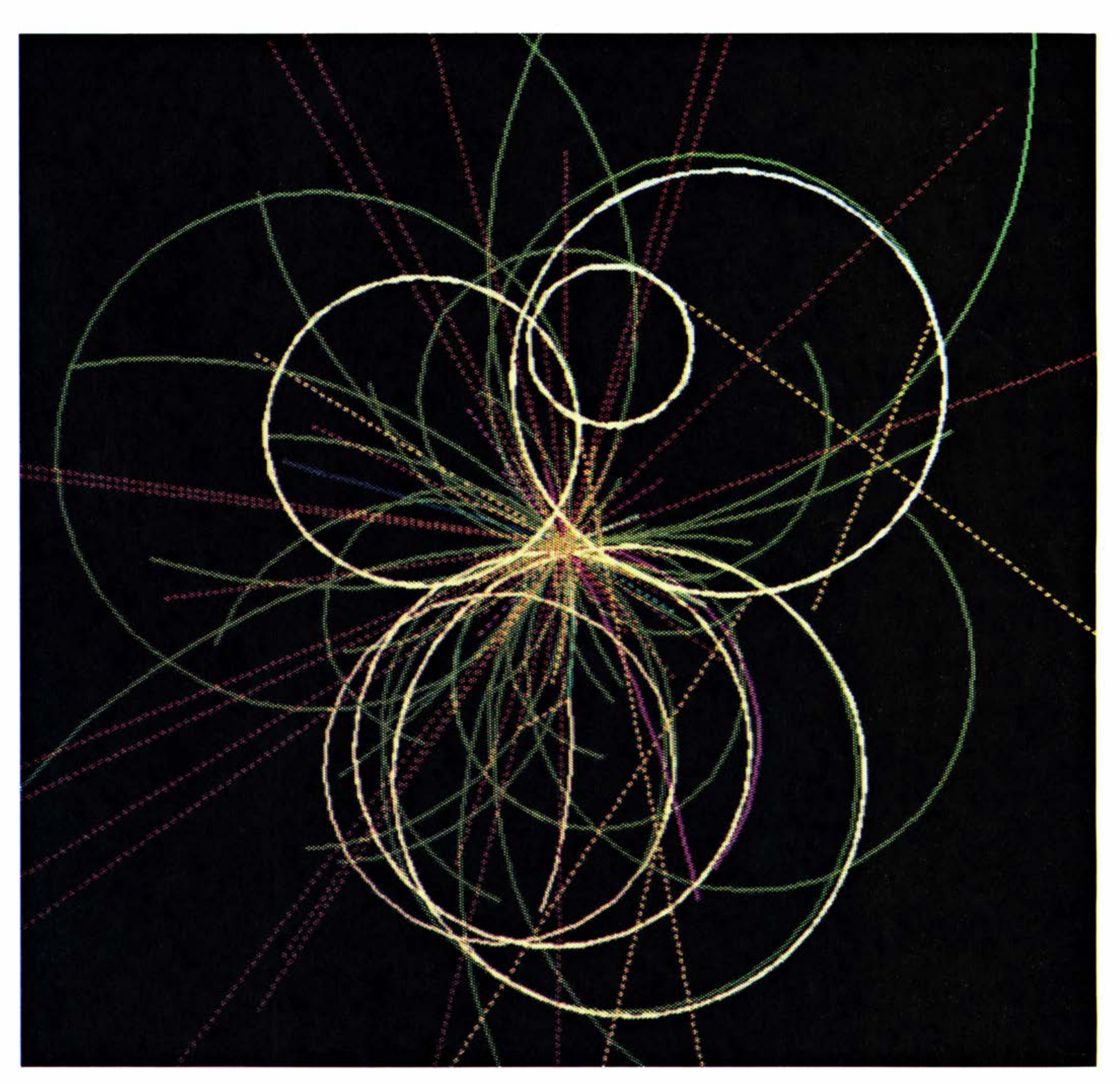
## Energía y luminosidad

La energía puede liberarse con suma eficiencia y crear nuevas partículas de gran masa si se provoca la colisión entre dos haces de momentos iguales y opuestos. Esta conclusión se deduce de la ley de conservación del momento, que se cumple estrictamente en todas las colisiones entre partículas: el momento total de los productos de una reacción debe ser igual al momento total de las partículas que reaccionan. Si una partícula de un haz energético choca contra una partícula de un blanco en reposo, el momento hacia adelante de la partícula energética entrante debe conservarse. Así, gran parte de la energía total de las dos partículas se destina a impartir momento hacia adelante a las partículas que abandonan el escenario de la colisión.

Por el contrario, si los momentos de las dos partículas colisionantes tienen la misma magnitud y direcciones opuestas, el momento total del par es cero. En principio, para comunicar momento a los productos de la reacción no

se requiere ninguna fracción de la energía total del par. La mayor parte de dicha energía está disponible para crear nuevas partículas. Nuestro diseño precisa ahora una máquina en la que dos haces de protones que giren en sentidos opuestos se aceleren hasta una energía de 20 billones de electronvolt (TeV) y choquen; la energía de una colisión sería de 40 TeV, superior en más de 20 veces a la energía de colisión del Tevatrón. La energía disponible para crear nueva materia en el SSC multiplicaría por 200 la energía disponible si uno solo de los haces se dirigiera contra un blanco fijo.

En un acelerador de haces colisionantes, la luminosidad del colisionador constituye un buen patrón para medir la intensidad útil de los haces. Formados éstos por trenes regularmente espaciados de paquetes iguales de partículas, la luminosidad es el número de partículas por paquete en un haz, multiplicado por el número de partículas interceptadas por unidad de área en el segundo haz. Y, todo ello, multiplicado por la frecuencia con la que los paquetes chocan. La luminosidad indica, pues, el número medio de sucesos por



**1. COLISION SIMULADA de dos protones, acelerado, cada uno, a una energía de 20 billones de electronvolt (TeV); da lugar a la lluvia de partículas que se muestra en esta imagen generada por ordenador. Las energías superan en un factor de 20 la energía que puede alcanzarse con los aceleradores de partículas más potentes que hay hoy en funcionamiento. Podrán alcanzarse, no obstante, en el acelerador propuesto, el supercolisionador superconductor (SSC). En la simulación por ordenador, una nueva partícula, llamada bosón de Higgs, cuya existencia está prevista por la teoría aunque no ha recibido todavía confirmación experimental, se materializa a partir de la energía liberada en el lugar de la colisión entre los protones. En la simulación, el bosón de Higgs tiene una masa de 300.000 millones de electronvolt (GeV); se desintegra en un bosón $W^+$ y otro $W^-$, bosones que transportan la fuerza débil responsable de la desintegración beta. Ni el bosón de Higgs ni los bosones $W^+$ y $W^-$ viven lo suficiente para ser detectados; su realidad se infiere de sus productos de desintegración. El bosón $W^+$ se desintegra en un positrón (*línea continua azul*) y un neutrino (*línea a trazos amarilla dirigida desde el centro hacia la parte superior izquierda*), mientras que el bosón $W^-$ se desintegra en dos quarks ligeros, cada uno de los cuales desencadena un chorro de partículas compuestas, formado principalmente por piones (*líneas continuas verdes*). Otros productos de la colisión se han simbolizado con otros colores: fotones en rojo, muones en blanco y bariones en púrpura. La simulación fue realizada por James Freeman, del Fermilab, utilizando el modelo ISAJET, desarrollado por Frank E. Paige, Jr., del laboratorio de Brookhaven.**

segundo en una reacción dada. Se necesita una alta luminosidad para garantizar que las reacciones de alta energía extraordinariamente raras se generen con suficiente frecuencia, y así confirmar su existencia a los físicos. El diseño actual del SSC necesita una luminosidad de $10^{33}$ unidades estándar, unas 1000 veces mayor que la proyectada para el Tevatrón.

Durante el desarrollo de este diseño, los autores han sido muy conscientes de la necesidad de una optimización económica. El SSC fue recomendado, en 1983, por el Comité Consultivo de Física de Altas Energías del Departamento de Energía de los EEUU, tras una dolorosa decisión que implicaba el abandono de otras varias peticiones de financiación de aceleradores de partículas. En respuesta a la recomendación, el Departamento de Energía encargó un programa de investigación y desarrollo cuyo propósito era preparar un proyecto que se moviera con realismo en lo técnico y ahorrador en lo económico, al objeto de elaborar una estimación detallada del coste. El programa, coordinado por el Grupo Central de Diseño del SSC de la Asociación de Universidades para la Investigación, se había de desarrollar en el Laboratorio Nacional de Brookhaven, en el Fermilab, en el Laboratorio Lawrence de Berkeley, en el Acelerador de Texas, en universidades y en la industria, con importante ayuda del extranjero. Una vez presentado, el informe del programa pesará en la decisión sobre el posterior apoyo del Departamento de Energía al proyecto SSC.

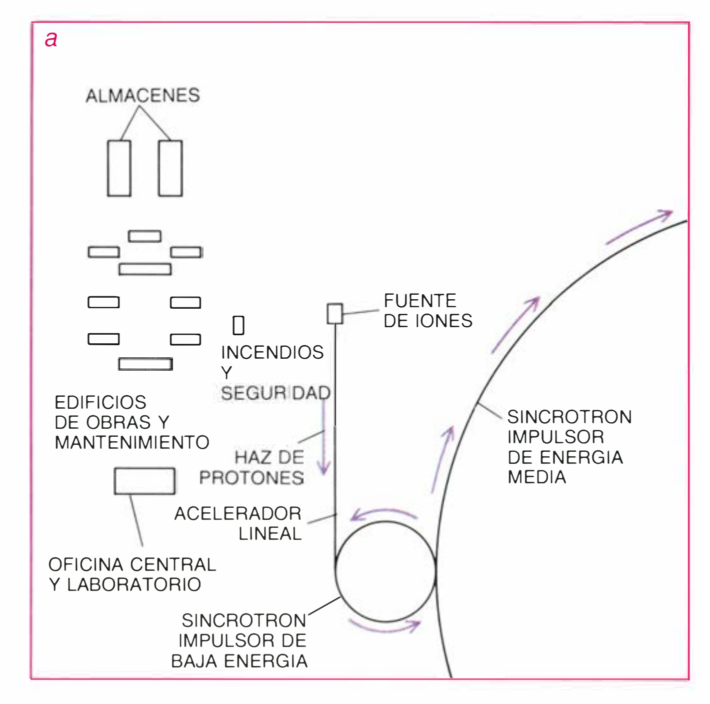
Ulteriores desarrollos ingenieriles del SSC aliviarán aún más su coste. No obstante, y a pesar de los esfuerzos, el coste final será alto en comparación con lo invertido en otras instalaciones científicas. El coste está marcado principalmente por una de las leyes básicas de la naturaleza: el examen de la estructura de la materia a resoluciones progresivamente mayores exige la aceleración de las partículas a energías de colisión cada vez más altas. Para justificar la inversión, se intentará explicar por qué los especialistas en partículas creen que vale la pena explorar las condiciones físicas extremas que hará posibles el SSC y cómo tales condiciones se materializarán con el proyecto presentado.

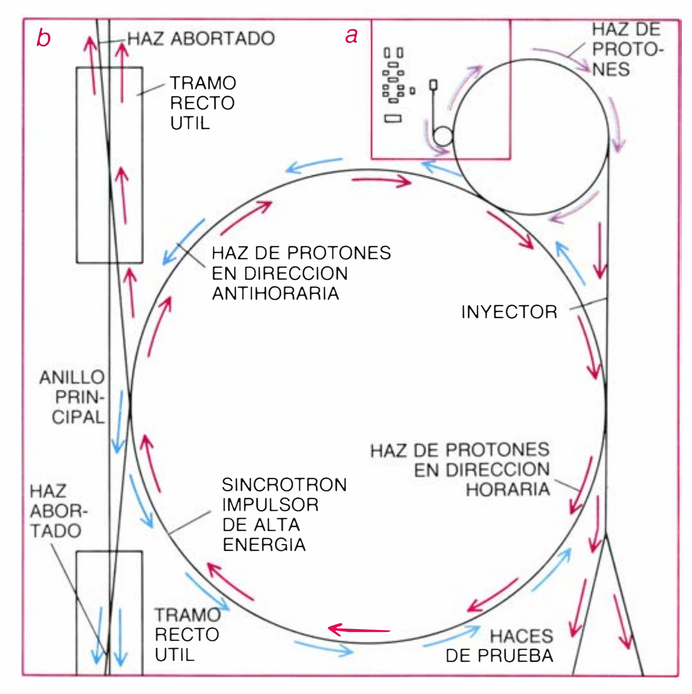
## El modelo estándar

A lo largo de los últimos veinte años se han realizado notables progresos en el descubrimiento de los constituyentes fundamentales de la materia y las fuerzas a través de las cuales interaccionan. De acuerdo con el modelo estándar de los procesos elementales, la materia está formada por quarks y leptones, con interacciones mutuas mediadas por el intercambio de las llamadas partículas de aforo ("gauge particles"). Se cree también que existen cuatro clases de interacciones: electromagnética, débil, fuerte y gravitatoria.

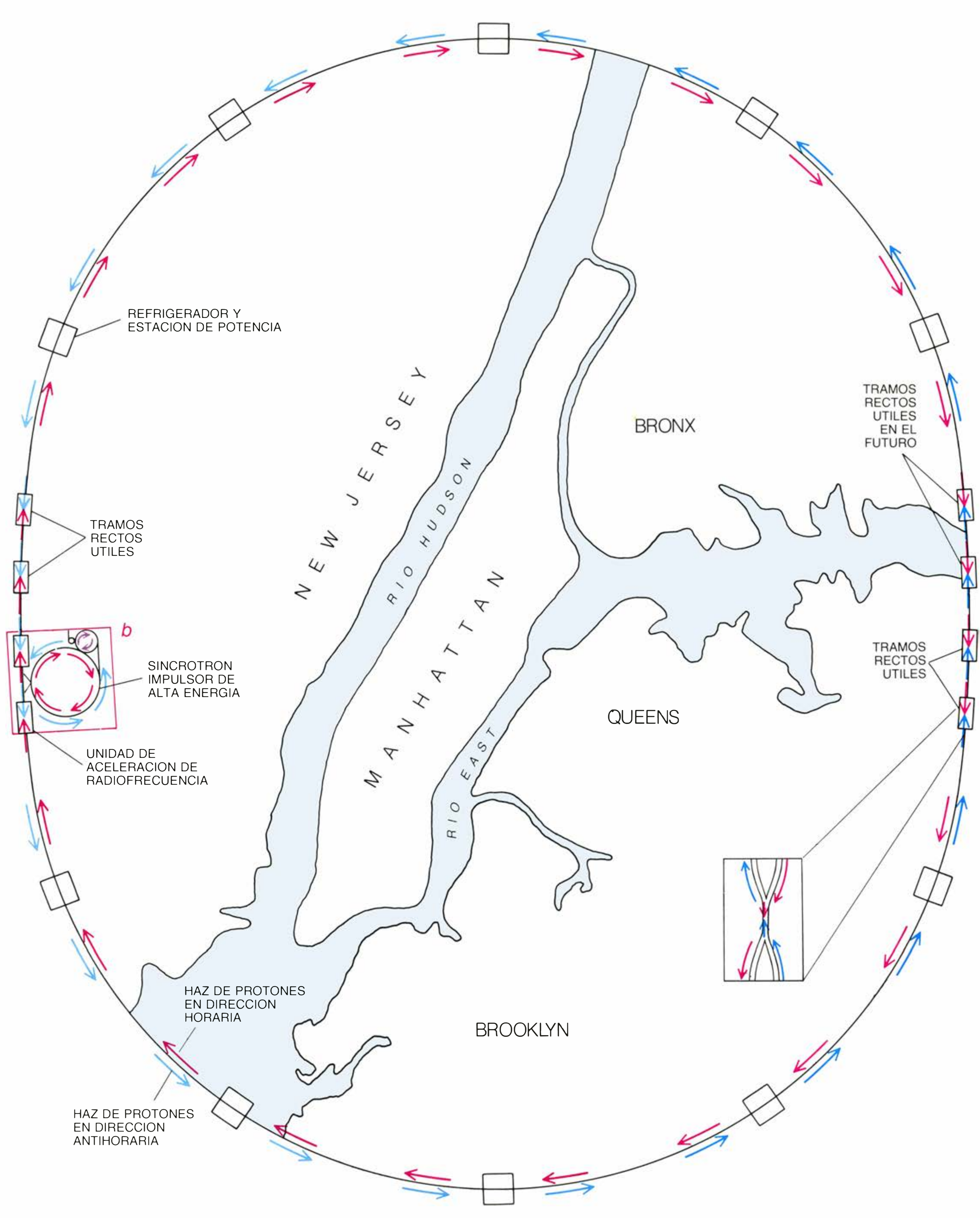
Por ejemplo, el electrón se clasifica entre los leptones y una partícula de aforo llamada fotón media sus interacciones electromagnéticas con el protón. La desintegración beta, fundamental en la combustión nuclear del Sol, es consecuencia de la interacción débil y está mediada por el intercambio de partículas de aforo llamadas bosones vectoriales débiles. El protón, el neutrón y muchas otras partículas se agrupan entre los hadrones. Están formados por tres quarks de carga fraccionaria. Los quarks se mantienen unidos mediante una interacción fuerte –llamada interacción de color–, que está mediada por el intercambio de ocho clases de partículas de aforo (gluones). Por analogía con estas tres interacciones, se supone que otra partícula de aforo, el gravitón, hace de mediadora en la interacción gravitatoria; pero no se ha detectado todavía. En conjunto, se cree que hay, por lo menos, seis quarks, con sus seis antiquarks correspondientes, cada uno de los cuales ofrece tres variedades de "color", seis leptones y sus seis correspondientes antipartículas, un fotón, tres bosones vectoriales débiles, ocho gluones y, quizás, un gravitón.

El modelo estándar se basa en datos de los grandes sincrotrones de protones del Fermilab y del CERN, el laboratorio

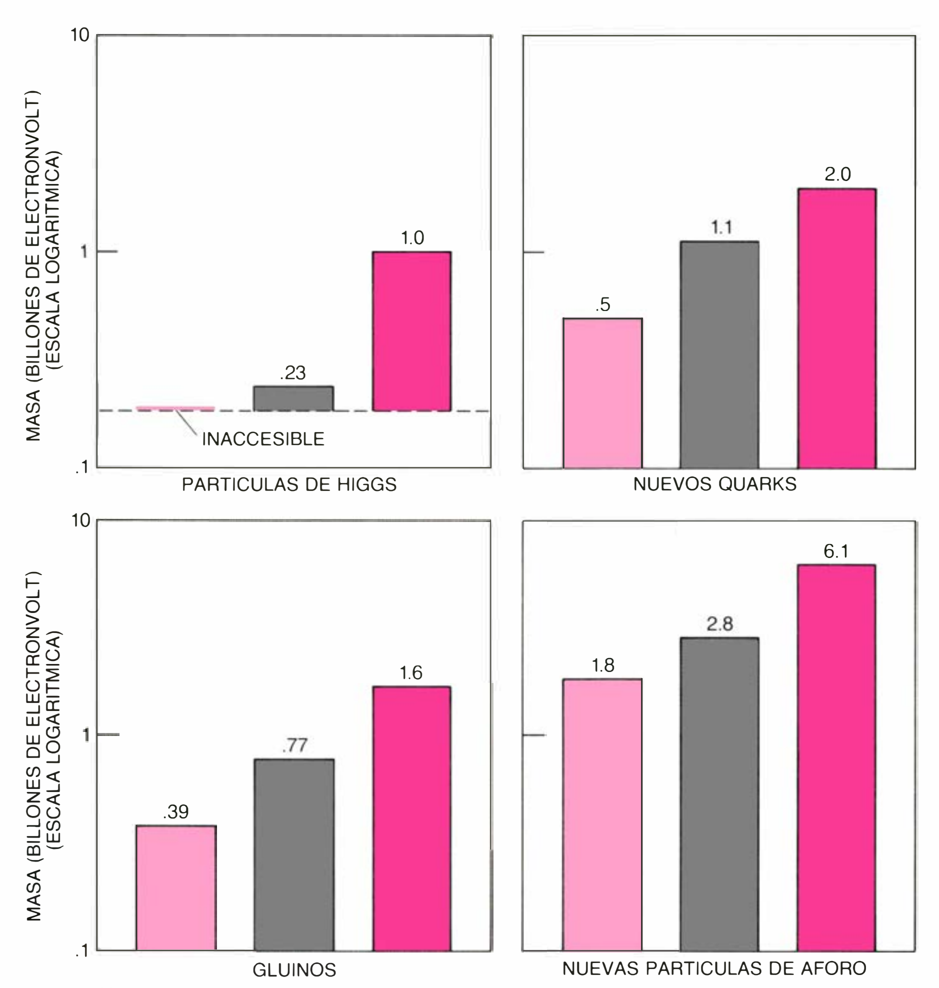



**2. PLANOS ESQUEMATICOS del supercolisionador superconductor (SSC). Representan la máquina a tres escalas: la oficina central y el complejo del laboratorio (*a*), la serie sucesiva de tres sincrotrones que preaceleran los protones (*b*) y el anillo principal del colisionador, superpuesto a un esquema del área metropolitana de la ciudad de Nueva York, dibujada a la misma escala (*c*). Los átomos de hidrógeno se ionizan y los protones que forman sus núcleos se aceleran (*púrpura*) en un acelerador lineal, hasta una energía de 0,6 GeV. Entonces, los protones entran en los sincrotrones y se aceleran hasta una energía de 1 TeV en**




**tres etapas; el haz se envía, alternativamente, en el sentido de las agujas del reloj y en sentido antihorario (*rojo, azul*) alrededor del sincrotrón impulsor de alta energía para inyectarse en los anillos del colisionador principal. En el anillo principal, cada haz se acelera hasta 20 TeV. Se provoca ahora el choque entre protones. Se controlan los productos de la colisión (*véase el recuadro de c*). Los imanes superconductores se enfrían mediante helio líquido rodeado de nitrógeno líquido, ambos bombeados a lo largo del anillo principal, desde diez refrigeradores y estaciones de potencia. (Ilustración recreada por Gabor Kiss.)**




**3. POR LIMITE DE DESCUBRIMIENTO entendemos la masa máxima que puede adquirir una partícula hipotética que haya de observarse merced a un acelerador de haces colisionantes de una energía de haz y una luminosidad dadas. Los gráficos indican que los límites de descubrimiento dependen de la energía del haz y de la luminosidad del acelerador, y esto para cuatro clases de partículas hipotéticas. Muestran también el límite inferior de la masa del bosón de Higgs: 0,18 TeV. Un descubrimiento se define como la creación de diez o más sucesos unívocamente identificados durante un año de toma de datos. El gráfico de color suave muestra, para cada partícula, el límite de descubrimiento en un colisionador protón-antiprotón cuya energía de haz fuera 6 TeV y cuya luminosidad fuera de 3 × $10^{31}$ unidades estándar. Tal colisionador podría construirse con los medios técnicos actuales en el túnel de 27 kilómetros del gran anillo electrón-positrón del CERN. Las columnas en gris simbolizan los límites de descubrimiento de un colisionador protón-antiprotón de 20 TeV cuya luminosidad fuera de 3 × $10^{31}$. Las columnas en color fuerte indican los límites de descubrimiento para el SSC: un colisionador protón-protón de 20 TeV cuya luminosidad sería de $10^{33}$.**

europeo de física de partículas, del Acelerador Lineal de Stanford (SLAC), de los colisionadores electrón-positrón de Cornell, Hamburgo y Stanford y, recientemente, del colisionador protón-antiprotón del CERN. Experimentos refinados de baja energía han aportado también importantes contribuciones a su conocimiento. Introducidos en un comienzo como meras entidades teóricas, cuando ya se habían descubierto centenares de hadrones, los quarks venían a restablecer alguna organización subyacente bajo las partículas "elementales" proliferantes. Los quarks adquirieron visos de realidad con los resultados de múltiples experimentos. Sin embargo, hubo que esperar a 1974 para respaldar sólidamente la creencia puesta en su existencia: se produjo el descubrimiento simultáneo en el SLAC y en Brookhaven de la partícula *J*/psi, predicha por la hipótesis de los quarks.

Un componente central del modelo estándar es la teoría electrodébil. En las actuales versiones de dicha teoría, los seis quarks y los seis leptones se agrupan en tres generaciones; a cada una de éstas se asigna un par de quarks y un par de leptones. Las interacciones electromagnética y débil serían aspectos diferentes de una interacción subyacente llamada electrodébil. La teoría electrodébil dicta predicciones precisas acerca de múltiples fenómenos y ha recibido confirmación cabal en muchos experimentos. Su confirmación más espectacular tuvo lugar en el CERN, en 1983, con la detección de los tres bosones vectoriales débiles, las partículas $W^+$, $W^-$ y $Z^0$.

La teoría electrodébil se inscribe en una tradición del pensamiento científico que arranca de la Grecia clásica: la unificación de diversos fenómenos en un mismo núcleo de nociones. De hecho, para muchos físicos, es el paradigma de la clase de teoría que un día puede lograr una explicación unificada de las cuatro interacciones fundamentales de la naturaleza. Según la teoría electrodébil, la unificación de las interacciones débiles y electromagnéticas se manifiesta sólo a energías extremadamente altas. A dichas energías, las interacciones son equivalentes porque las masas de los bosones de aforo que median las dos interacciones son despreciables. La plena simetría de las dos interacciones puede patentizarse sin inhibición.

La hipótesis de tal simetría a altas energías contrasta fuertemente con las propiedades de las dos interacciones en las condiciones ordinarias de los laboratorios. Allí, el alcance de la interacción débil es aproximadamente mil veces menor que el diámetro del núcleo atómico, mientras que el alcance de la interacción electromagnética es infinito. Según la teoría electrodébil, esta diferencia débese a que los bosones de aforo débiles son partículas muy pesadas, mientras que la masa del bosón de aforo electromagnético (el fotón) es nula. Se dice que la simetría de las dos interacciones está rota.

## El origen de la masa

¿Por qué se rompe a energías ordinarias la simetría de las interacciones débiles y electromagnéticas? La pregunta se hace aún más apremiante si recordamos que, en la formulación teórica de la simetría electrodébil, tanto el fotón como los bosones vectoriales débiles presentan inicialmente masa cero. Así, las grandes masas observadas para los bosones vectoriales débiles surgen de la rotura de la simetría electrodébil y la pregunta anterior se transforma en esta otra: ¿Cuál es el origen de la masa?

El origen de la masa es un problema central cuando se trata de determinar las posibilidades fundamentales del SSC. El análisis de la rotura de la simetría electrodébil conduce a varios escenarios teóricos posibles, aunque todos comparten una propiedad común: la prueba a favor, o en contra, de cualquiera de ellos debe ponerse de manifiesto con las energías de colisión que ahora se proponen para el SSC. Hay, sin duda, muchas otras cuestiones abiertas dentro del modelo estándar o fuera del mismo, que serán objeto de análisis con el SSC. Y es casi seguro que existen procesos físicos completamente desco-

nocidos que el SSC sacará a la luz. No obstante, la mayoría de las indicaciones teóricas, relativas a la escala adecuada del SSC, que puedan obtenerse, vendrán, sobre todo, de la capacidad de abordar el origen de la masa.

En la versión más sencilla de la dinámica electrodébil, la rotura espontánea de simetría se debe a un campo eléctricamente neutro, llamado campo de Higgs en honor de Peter W. Higgs, de la Universidad de Edimburgo. El campo de Higgs, si existe, poseerá un valor de fondo uniforme y no nulo, incluso en el vacío. La idea de que el vacío "contiene" algo, incluso un campo uniforme no nulo, choca contra la noción intuitiva de un vacío como espacio sin nada. Sin embargo, en mecánica cuántica, el aspecto paradójico de tal resultado se ha disipado desde hace tiempo. El vacío mecánico-cuántico fluctúa activa y constantemente, sea o no real el campo de Higgs.

La interacción entre partícula y campo de Higgs contribuye a la energía de la partícula con respecto al vacío. Tal energía es equivalente a una masa. En el más sencillo de los modelos del campo de Higgs, las masas de los quarks, los leptones y los bosones vectoriales débiles se explican en virtud de la interacción con un único campo de Higgs. Existe siempre una partícula asociada a un campo mecanico-cuántico; así, en la forma más sencilla del mecanismo de Higgs para la rotura de simetría, hay una partícula de Higgs asociada a un campo de Higgs. Si la partícula de Higgs existe, debería detectarse; hasta la fecha, sin embargo, la búsqueda de la misma ha resultado un fracaso.

## La partícula Higgs

Se plantea el problema siguiente: la teoría apenas si restringe el valor de la masa de la partícula de Higgs. Quizá tenga la ligereza de unos pocos miles de millones de electronvolt (GeV); o quizá llegue a pesar un teraelectronvolt. Si la masa no alcanza los 50 GeV, debería dejarse atrapar en los colisionadores electrón-positrón, tales como el colisionador lineal de Stanford (SLC) o el gran colisionador electrón-positrón (LEP), del CERN, que se supone estarán construidos y en funcionamiento hacia el final de la década actual. Si su masa está entre 50 y 200 GeV, el colisionador Tevatrón, del Fermilab, debería ser capaz de producirla, aunque establecer una prueba firme de su existencia, entre una muchedumbre de otras partículas en dicho intervalo de masas, puede convertirse en una hazaña experimental difícil.

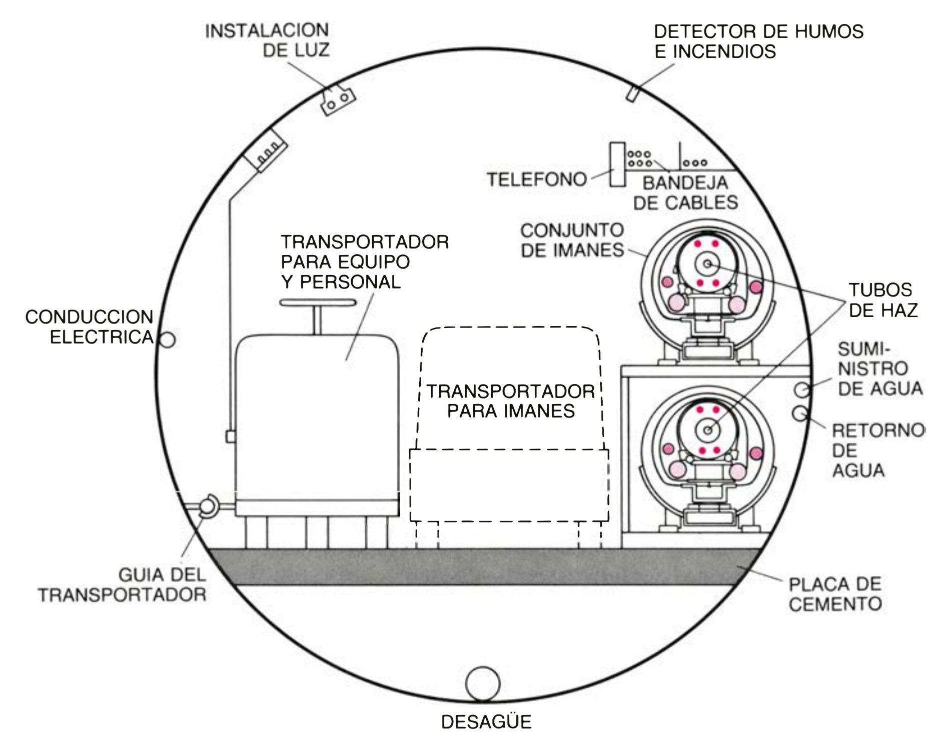
Para masas de Higgs muy por encima de los 200 GeV, no conviene considerar elemental la partícula de Higgs. La mecánica cuántica enseña que, cuanto más corta es la vida media de una partícula, menor certeza tenemos de su energía o, equivalentemente, de su masa. Si la masa de Higgs supera los 200 GeV, se desintegrará en dos $W$ o dos partículas $Z$ y su vida media será brevísima; hasta el punto que su masa quedará "difuminada" en un amplio intervalo. ¿Qué sentido tiene considerar partícula una entidad que carece de masa definida?

Si la masa de Higgs pesara 1 TeV, de acuerdo con lo predicho por la teoría electrodébil, a esas energías –o por encima de ellas– aparecerán fenómenos completamente nuevos. En tales circunstancias, la interacción electrodébil se torna fuerte. Como la dinámica electrodébil gobierna las interacciones entre leptones y quarks, tales partículas pueden combinarse, a energías de 1 TeV o más, en partículas compuestas de características nuevas y sorprendentes. En cualquier caso, para confirmar, o no, el mecanismo de Higgs analizando el posible intervalo de masas de una partícula de Higgs identificable, será necesario el SSC.

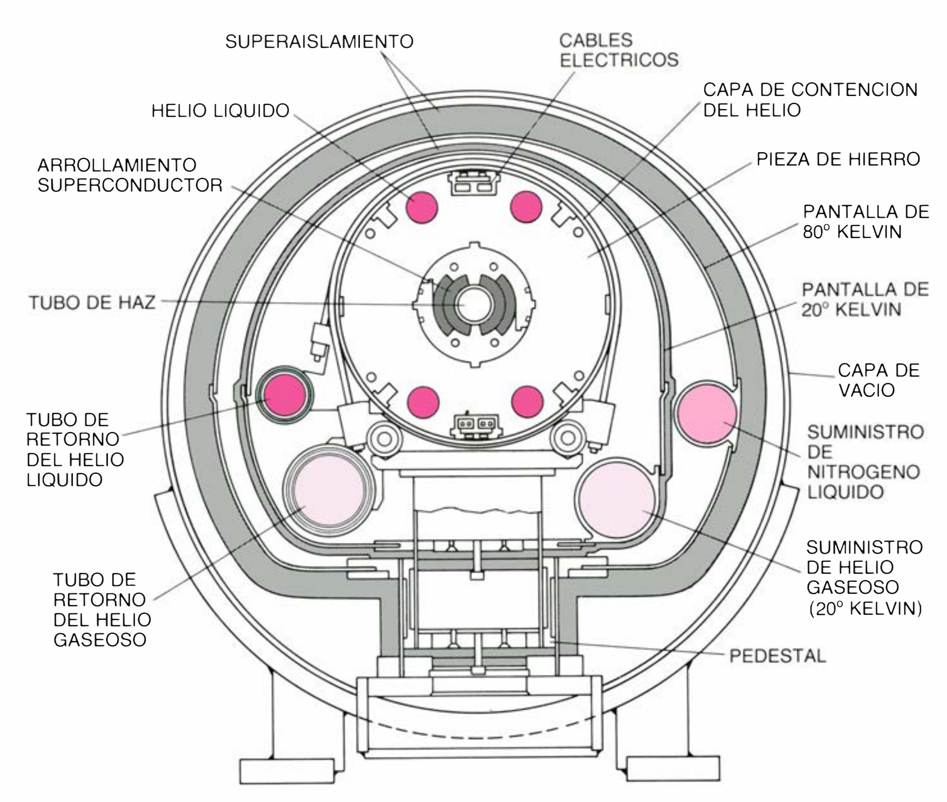
Muchos físicos opinan que la forma más sencilla del mecanismo de Higgs para la ruptura de simetría es sólo una aproximación de baja energía a la realidad. Se fundan en que la partícula de Higgs, si existe, no puede tener espín. Si lo poseyera, el campo de Higgs tendría también espín y la masa de una partícula ordinaria dependería de su orientación en el vacío. Tal dependencia rotacional no se ha observado.

En mecánica cuántica, el espín de una partícula sólo puede tomar valores discretos y las partículas que tienen espines enteros (0,1,2, etc.) se distinguen nítidamente de las partículas dotadas de espines semienteros (1/2, 3/2, etc.). Las partículas que tienen espín entero se llaman bosones; así, la partícula de Higgs, de espín cero, es un bosón, igual que los bosones de aforo observados, de espín 1, tales como el fotón, que median interacciones fundamentales. Las partículas con espín semientero reciben el nombre de fermiones; abarcan todos los quarks y leptones.

Para calcular la masa del bosón de Higgs, es obligado establecer ciertas hipótesis acerca de los procesos físicos a altas energías. Si el bosón de Higgs es una partícula elemental, la masa calculada varía ampliamente, incluso por pequeños cambios de tales hipótesis. Tal sensibilidad matemática no tiene una interpretación física natural; además, no es una característica de las expresiones de las masas de las partículas de espín 1/2. Por ello, algunos físicos teóricos han propuesto evitar la dificultad matemática y retener el necesario espín cero del bosón de Higgs suponiendo que el bosón sea una par-




**4. TUNEL PRINCIPAL propuesto para el SSC, en un corte transversal. Los dos tubos de haz para los dos haces de protones que giran en sentidos opuestos están a la derecha. El túnel mide tres metros de diámetro.**




**5. SECCION ESQUEMATICA DEL CONJUNTO DE IMANES que deberán montarse en el túnel del SSC. En la parte superior central del conjunto, uno de los haces de protones atraviesa un tubo en el que se ha hecho el vacío. El tubo está rodeado por arrollamientos de alambre superconductor; la corriente que pasa sin resistencia por el alambre crea el enorme campo magnético necesario para curvar el haz de protones. El resto del sistema que rodea al tubo del haz sirve para conservar el imán a las bajas temperaturas necesarias para mantener la superconductividad de los arrollamientos. La primera capa de conducción que rodea el imán transporta helio líquido refrigerante, mantenido a 4,35 grados Kelvin. Esta capa está envuelta por otra conducción que lleva nitrógeno líquido a 80 grados K. Capas de material aislante rodean la conducción.**

tícula compuesta, en vez de elemental, formada por dos fermiones de espín 1/2. Tales partículas compuestas abundan en otros dominios de la física: por ejemplo, el pion, de espín cero, es una partícula compuesta formada por dos quarks de espín 1/2. Los espines se anulan (cancelan) y forman un compuesto sin espín, porque los dos quarks giran en direcciones opuestas.

Un bosón de Higgs compuesto exigiría la existencia de una nueva familia entera de partículas pesadas de espín 1/2: los techniquarks. Tales partículas estarían sometidas a una nueva interacción fuerte –interacción de technicolor–, que se supone análoga a la interacción fuerte de color que liga los quarks formando hadrones. Los techniquarks no sólo se ligarían para crear el bosón de Higgs, sino que lo harían también para originar una plétora de otras technipartículas compuestas: technipiones, mesones technivectoriales, etcétera. Esas nuevas partículas serían bastante pesadas; algunas de ellas al menos presentarían, sin embargo, masas en el rango de los 50 a los 500 GeV. En el Tevatrón, y aceleradores parejos, el número generado de tales partículas sería pequeño, lo que entorpecería su detección sobre el fondo, aun cuando sus masas estuvieran cerca del límite inferior del intervalo esperado. Para comprobar la teoría, debería emplearse un colisionador, tal como el SSC, con una energía de haz superior a varios TeV.

## Supersimetría

Otro programa teórico que reviste características de especial atractivo es el de la supersimetría. Podría convertirse en la opción alternativa al mecanismo simple de Higgs para explicar el origen de la masa. En un mundo supersimétrico, cada partícula, incluido el bosón de Higgs, tiene una compañera idéntica en todo, salvo en su espín. A cada fermión ordinario le corresponde un bosón supersimétrico sin espín; por ejemplo, el electrón y el quark, de espín 1/2, poseen por compañeros respectivos el selectrón y el squark, de espín 0. A cada bosón ordinario le corresponde un fermión supersimétrico, de espín 1/2; por ejemplo: el compañero supersimétrico del fotón, de espín 1, es el fotino, de espín 1/2, el compañero del gluon, de espín 1, es el gluino, de espín 1/2, y el compañero del bosón de Higgs, de espín 0, es el higgsino, de espín 1/2.

Si las partículas supersimétricas existieran en la naturaleza, como copias exactas de sus compañeras menos en su espín, se habrían identificado la mayoría de ellas y no sólo una vez, sino muchas. Sin embargo, se han acometido búsquedas intensas sin el menor éxito: no hay pruebas de compañeras supersimétricas. Cabría suponer, pues, que el interés por la supersimetría hubiera ido apagándose. Pero hay razones para que no haya sido así. La primera: la existencia de partículas supersimétricas resolvería el problema de la sensibilidad matemática de la expresión teórica de la masa del bosón de Higgs. Una segunda: ¿es la supersimetría una simetría rota en nuestro mundo, como lo es la simetría electrodébil? Una supersimetría rota puede dar lugar a partículas supersimétricas sustancialmente más pesadas que sus compañeras ordinarias.

Nadie sabe cuán sensible es la masa a la rotura de la simetría. Por demostración experimental conocemos que la masa del selectrón, si existe, será, por lo menos, 40.000 veces mayor que la del electrón. ¿Implica esto que la supersimetría debe estar "muy" rota? No tenemos respuesta. Sí se sabe que, en el caso de que la supersimetría resulte correcta, introducirá también, como la teoría del technicolor, un nuevo mundo de partículas. La mayoría de ellas serían bastante masivas; si no lo fueran, ya se habrían detectado. Sin duda alguna, se requieren potentes y nuevos aceleradores para encontrarlas.

Hay una tercera razón, más general, para no abandonar la supersimetría o, en toda caso, cualquier otra teoría, como la del technicolor, con alguna posibilidad de explicar los misterios de la rotura de la simetría electrodébil y del origen de la masa. Con independencia de que el bosón de Higgs sea compuesto o elemental, esté o no incluido en una familia de partículas supersimétricas, o exista o no realmente, un principio general de la mecánica cuántica garantiza que nuevos fenómenos físicos, profundamente relacionados con el origen de la masa, comiencen a aparecer a energías cercanas a 1 TeV. Estos fenómenos son inevitables porque, si el modelo estándar en boga se extrapola sin correcciones a tal dominio energético, las probabilidades calculadas por la teoría para ciertas interacciones se hacen mayores que la unidad. Ahora bien, considerando que ninguna probabilidad verdadera puede ser mayor que 1, la teoría, tal como se nos ofrece, no puede estar completa.

Se desconoce la extensión teórica correcta del modelo estándar a energías

muy altas; *a fortiori*, no puede describirse aún la naturaleza exacta de los nuevos fenómenos físicos. Si el bosón de Higgs es bastante masivo, hemos señalado antes la posibilidad de que la interacción electrodébil se convierta en fuerte. Por otro lado, si el bosón de Higgs resulta ser ligero, su pequeña masa podría explicarse con la supersimetría. En dicho caso, el dominio de energías alrededor de los pocos TeV tendría abundancia de compañeras supersimétricas de las partículas conocidas. La posibilidad de analizar dicho dominio energético constituye, por tanto, un objetivo de extrema importancia para la compresión básica de la materia.

## Necesidad de 20 TeV

El technicolor y la supersimetría son teorías que predicen, de un modo específico, el límite de descubrimientos de un colisionador dotado de cierta energía y luminosidad; en otras palabras, acotan la máxima masa que puede tener una partícula hipotética, si debe crearse y detectarse en el colisionador. Podría pensarse que una máquina que produce colisiones de protones con una energía total de 40 TeV dejaría disponible aproximadamente esta cantidad para la creación de nuevas partículas. Desgraciadamente, sólo se libera una parte de la energía de la colisión. Todos los hadrones, así protones y antiprotones, son sistemas compuestos, parecidos a saquitos de canicas. La energía total se divide entre los quarks, antiquarks y gluones que forman cada hadrón, con lo que una colisión sólo puede liberar la cantidad de energía que lleva cualquier par de constituyentes colisionantes. Por ejemplo, el Tevatrón, con una energía total de 1,8 TeV, sólo puede explorar un rango de masas que llega hasta unos 0,3 TeV.

Para dar una idea de cómo la máxima masa detectable varía con el diseño del colisionador, consideremos la posible existencia del gluino y del squark. Para un colisionador protón-antiprotón, cada uno de cuyos haces se acelera a una energía de 6 TeV, y cuya luminosidad fuera de $3 \times 10^{31}$ unidades estándar, el gluino, o el squark, más pesado detectable tendría una masa de unos 0,4 TeV. Con los medios técnicos disponibles, tal colisionador podría construirse en el túnel del LEP, en el CERN. En el SSC, donde la luminosidad aumentará respecto a tal colisionador LEP en un factor de 30, y la energía se multiplicará en un factor de tres, el gluino, o el squark, más pesado detectable tendría una masa de unos 1,5 TeV. De manera similar, el SSC podría detectar bosones de Higgs de 1 TeV de masa, nuevos quarks de 2 TeV y nuevas partículas de aforo de hasta 6 TeV de masa [*véase la figura 3*].

Los pormenores de las numerosas estimaciones teóricas y de las hipótesis sobre la máquina admiten objeciones y discrepancias, pero el mensaje es claro. Para analizar con eficacia el rango de masas de partículas del orden de 1 TeV, un colisionador con las características diseñadas para el SSC en proyecto es el adecuado. Menor energía o menor luminosidad empezarían a comprometer los posibles descubrimientos. Además, si la máxima energía disponible estuviera justo debajo del umbral de alguna física radicalmente nueva, ésta quedaría sin descubrir, con independencia de cuál fuera la luminosidad. La necesidad de las máximas energías de haz posibles resulta primordial.

## La trayectoria del protón

Para comprender cómo el SSC logrará la capacidad y posibilidades previstas, nos será útil seguir la trayectoria de los protones a lo largo del colisionador, desde su fuente hasta los lugares de choque [*véase la figura 2*]. Los protones empiezan su viaje como núcleos de átomos de hidrógeno ionizados en un gas. Se extraen del gas mediante electrodos convenientemente dispuestos y emergen con una energía cinética de unos pocos miles de electronvolt.

Desde allí, penetran en un acelerador lineal una serie de electrodos que aceleran los protones a través de un conjunto de pequeños impulsos. Los protones son arrastrados por una onda de diferencias de potencial, sincronizada con precisión, que se mueve entre los electrodos. La aceleración eleva la energía del haz hasta 600 millones de electronvolt. El haz entra, entonces, en el primero de una secuencia de cuatro anillos de sincrotrón.

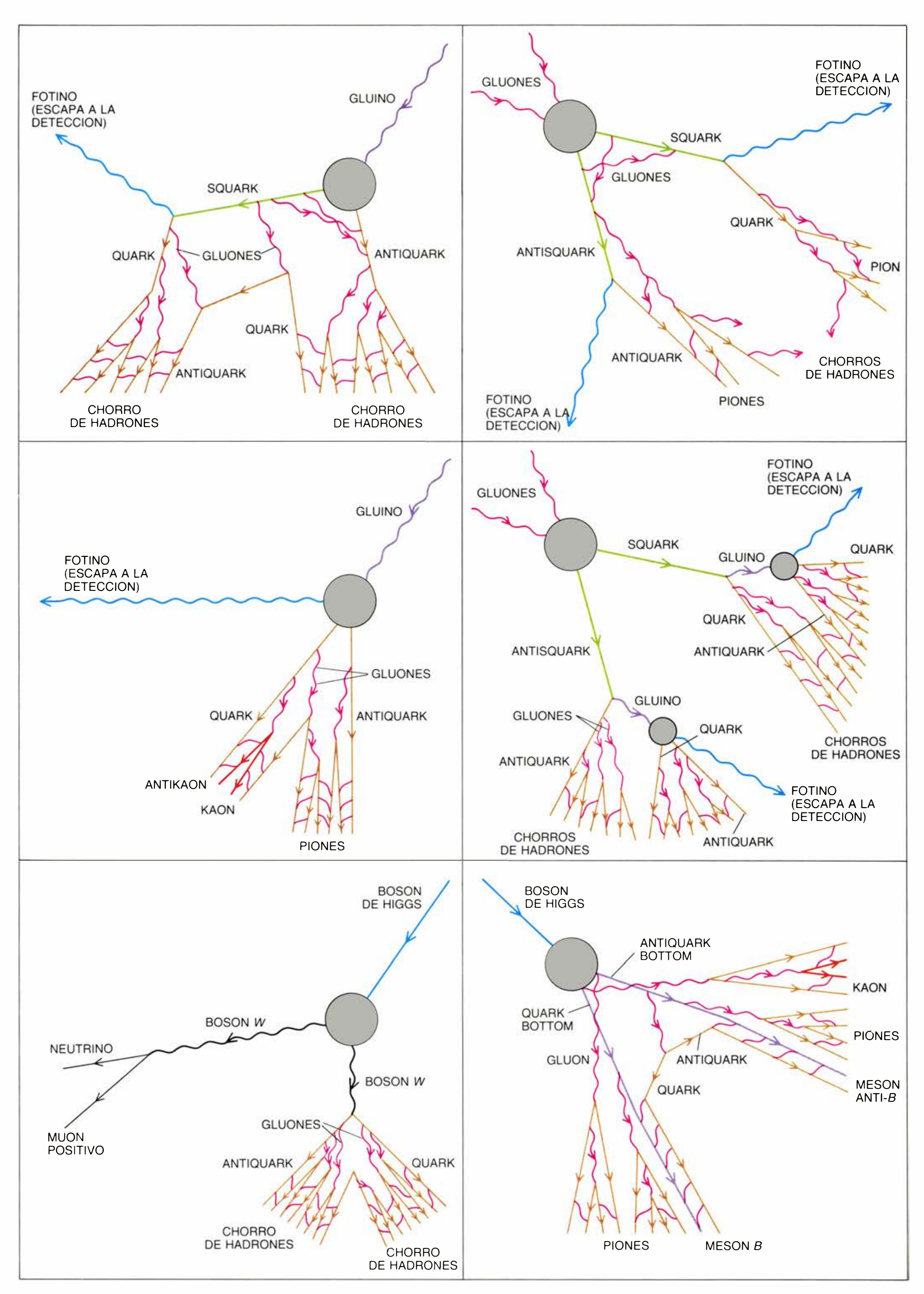
En la aceleración sincrotrónica, un campo magnético uniforme obliga a los protones a seguir una trayectoria predeterminada. En una parte especial del anillo del sincrotrón, el haz atraviesa un acelerador lineal y, por tanto, su energía aumenta a cada vuelta al anillo. A medida que crece la energía, aumenta también la intensidad del campo magnético a fin de mantener los protones en su órbita cerrada. La frecuencia de la onda viajera que acelera los protones en el acelerador lineal debe estar sincronizada con la frecuencia con la que los protones dan vueltas al anillo; la frecuencia está en la zona de radio del espectro electromagnético. El alto rendimiento de los sistemas de aceleración modernos facilita que las zonas de aceleración ocupen sólo una longitud de 23 metros en la circunferencia completa, de 85 kilómetros, del SSC.

El coste de los imanes que dirigen el haz es proporcional al diámetro de la región del tubo del haz en la que se necesita un campo magnético uniforme. El diámetro del campo depende del diámetro del haz; éste, a su vez, depende, entre otras cosas, de la razón entre el momento de las partículas en una dirección transversal al haz y su momento a lo largo de la trayectoria del haz. Dicha razón decrece continuamente porque el sistema de aceleración aumenta el momento de las partículas sólo en la dirección del haz. En un gran sincrotrón, la disminución del diámetro del haz se utiliza para minimizar el diámetro del campo magnético uniforme necesario en cada paso de la aceleración, abaratando así el coste total. En el diseño actual, una secuencia de sincrotrones dirige los protones a través de una serie de tubos de haz progresivamente más estrechos. El primer sincrotrón acelerará los protones desde una energía de 0,6 GeV hasta unos 8 GeV; el segundo los empujará desde 8 a unos 100 GeV y el tercero los llevará desde 100 GeV hasta 1 TeV. El anillo sincrotrónico principal del SSC, el mayor, acelerará los protones hasta una energía final de 20 TeV.

A la energía máxima, el diámetro del haz medirá un quinto de milímetro. A fin de incrementar la luminosidad del colisionador, los haces se focalizarán entonces hacia los puntos de colisión mediante potentes lentes magnéticas, formando paquetes más apretados todavía. Tras la focalización, los paquetes se habrán convertido en finísimos cilindros de unos diez micrometros de diámetro por 15 centímetros de longitud. Cada paquete contendrá unos 10.000 millones de protones; la densidad de cada uno será, pues, una diezmilésima parte de la densidad de las moléculas en el aire, a presión y temperaturas ordinarias. Como la probabilidad de colisión entre los protones de dos paquetes que avanzan girando en sentidos opuestos es pequeña, los paquetes pueden interpenetrarse repetidamente en los puntos de interacción durante muchas horas, sin necesidad de rellenarlos.

## Imanes superconductores

La necesidad de curvar y focalizar el haz de protones convierte al sistema de imanes en un elemento clave de cualquier sincrotrón. En los dos últimos años, los Estados Unidos han puesto



FOTINO (ESCAPA A LA DETECCION)
GLUINO
SQUARK
QUARK
GLUONES
ANTIQUARK
QUARK
ANTIQUARK
CHORRO DE HADRONES
CHORRO DE HADRONES
GLUONES
FOTINO (ESCAPA A LA DETECCION)
SQUARK
GLUONES
QUARK
ANTISQUARK
PION
ANTIQUARK
CHORROS DE HADRONES
PIONES
FOTINO (ESCAPA A LA DETECCION)
GLUINO
FOTINO (ESCAPA A LA DETECCION)
GLUONES
QUARK
ANTIQUARK
ANTIKAON
KAON
PIONES
GLUONES
FOTINO (ESCAPA A LA DETECCION)
SQUARK
GLUINO
QUARK
QUARK
ANTISQUARK
ANTIQUARK
GLUINO
GLUONES
QUARK
CHORROS DE HADRONES
ANTIQUARK
FOTINO (ESCAPA A LA DETECCION)
CHORROS DE HADRONES
ANTIQUARK
BOSON DE HIGGS
BOSON W
NEUTRINO
BOSON W
MUON POSITIVO
GLUONES
ANTIQUARK
QUARK
CHORRO DE HADRONES
CHORRO DE HADRONES
BOSON DE HIGGS
ANTIQUARK BOTTOM
KAON
QUARK BOTTOM
PIONES
GLUON
ANTIQUARK
MESON ANTI-B
QUARK
PIONES
MESON B

especial empeño en el estudio de diferentes clases de imanes que podrían adoptarse en el SSC. En principio, sería posible construir el SSC con electroimanes de conductor de cobre, tales como los que se utilizaban en los sincrotrones de los años 50 y 60. Tal anillo, una vez fabricado para acelerar protones a una energía de 20 TeV, consumiría, por lo menos, 4000 millones de watt de potencia y elevaría los costes de funcionamiento a unas cifras astronómicas. Además, las propiedades magnéticas del hierro y la capacidad del cobre para conducir la corriente eléctrica limitan los campos electromagnéticos ordinarios a una intensidad de unos 2 teslas, o sea, 20.000 gauss, que viene a equivaler a los campos generados en los motores eléctricos de los aparatos domésticos.

Los imanes superconductores resuelven, a la baja, ambos problemas. Reducen drásticamente la potencia total consumida por el sistema acelerador y crean campos magnéticos varias veces más intensos que los generados con imanes convencionales. Un campo magnético más intenso permite confinar los protones de una energía dada en órbitas más apretadas, reduciendo el tamaño necesario del anillo del sincrotrón.

En un electroimán convencional, arrollado con alambre de cobre, la potencia se disipa generando corrientes eléctricas a través del alambre. Sin embargo, cuando ciertos metales, aleaciones y compuestos metálicos, se enfrían por debajo de determinada temperatura crítica, se hacen superconductores: conducen la corriente eléctrica sin resistencia. En un imán arrollado con alambre superconductor, la única potencia exigida es la necesaria para mantener el alambre por debajo de la temperatura crítica de superconducción. Los refrigeradores necesarios para el SSC consumirán unos 30 millones de watt de potencia, cifra ligeramente inferior al total necesario en el mayor complejo acelerador actual.

En 1983 entró en servicio el primer sincrotrón superconductor, el Tevatrón. Desde entonces ha funcionado con pleno éxito, suministrando datos de sumo interés para el proyecto del SSC. En septiembre del año pasado, los autores y sus colegas del grupo central de diseño del SSC cerraron una etapa al optar por el diseño de los imanes superconductores de desviación. Los imanes desviadores tendrán una intensidad de campo de unos 6,6 teslas, y esta intensidad exigirá una circunferencia de casi 84 kilómetros para el anillo sincrotrónico principal.

Tanto el arrollamiento superconductor del imán como la pieza de hierro circundante se mantendrán a una temperatura de 4,35 grados Kelvin mediante un flujo presurizado de helio líquido. Diez refrigeradores, regularmente espaciados a lo largo del anillo, aportarán, cada uno, una corriente refrigerante de unos cuatro kilómetros en ambos sentidos. Una pantalla térmica intermedia de nitrógeno líquido, mantenido a 80° K, reducirá notablemente la energía calorífica incidente sobre el helio líquido. Para estabilizar el sistema entero frente a los cambios de temperatura y las vibraciones mecánicas, el anillo se introducirá en un túnel. Este se cubrirá luego con seis metros de tierra, por lo menos, a fin de absorber la radiación ionizante que pudiera generarse si el haz incidiera en las paredes del tubo del haz.

## Aparatos de detección

Como el SSC explorará un intervalo de energías hasta ahora desconocido, nadie puede predecir con certeza las propiedades de los sucesos más interesantes. Resulta, pues, imprescindible maximizar la variedad y flexibilidad de los detectores que se proyecten. Los ensayos con el sistema inicial de detectores se encargarán de guiar el programa experimental definitivo.

Se provocarán las colisiones entre los haces de protones que giran en sentidos opuestos obligándolos a cruzarse frontalmente, o bajo un ángulo muy pequeño, en uno o varios puntos del anillo. En tales puntos de intersección habrá refinados detectores, colocados en salas experimentales. Los detectores registrarán el paso de partículas que surjan de las colisiones y las rastrearán electrónicamente a fin de identificar las que presenten interés potencial.

Esperamos generar hasta 100 millones de colisiones por segundo en cada región de interacción del SSC; el escrutinio electrónico no es, en absoluto, trivial: debe encontrarse una aguja o, por lo menos unas pocas agujas, en un pajar de datos. Además, una vez que un suceso de colisión potencialmente interesante quede identificado, debe someterse a pruebas adicionales más detalladas. Si consigue pasar las pruebas, debe recogerse y registrarse suficiente información acerca de las trazas que salen de la colisión de suerte que permita una posterior reconstrucción de lo que realmente ha sucedido. El proceso entero debe hacerse deprisa, al objeto de minimizar el tiempo "muerto" en el que el detector y su sistema de computación asociado no están recogiendo datos de colisiones.

## Detección directa e indirecta

Como la necesidad de evitar tiempo muerto exige que se registre sólo una pequeña fracción de todas las colisiones, debe diseñarse un disparador electrónico para los fenómenos interesantes. Aquí sirven de guía tanto la historia como las ideas. La emisión de electrones y muones energéticos señaló, recuérdese, la presencia de los bosones $W^+$, $W^-$ y $Z^0$ y, en el caso de los dos bosones $W$, los señaló también la emisión de neutrinos. Se espera que el bosón de Higgs dé lugar a señales de desintegración similares a las de los bosones débiles. La teoría predice que debe desintegrarse en partículas de la mayor masa posible, tales como quarks pesados o bosones $W$. A su vez, los quarks y bosones $W$ deben dar lugar a electrones, muones y neutrinos. Ni el bosón de Higgs, ni los quarks pesados, ni los bosones $W$ viven lo suficiente para detectarse directamente, de manera que su presencia debe inferirse a partir de sus subproductos.

La presencia de electrones y muones, cargados eléctricamente, admite una detección directa. Más sutil resulta, sin embargo, la detección de neutrinos. La presencia del neutrino puede inferirse midiendo el momento total, en direcciones transversales a los haces de protones, de las partículas que emergen de una colisión. Como el momento total de los protones colisionantes en las direcciones transversales a los haces de protones es básicamente nulo, el momento transversal total de las partículas emergentes debe ser también cero. Así, el momento total de las partículas en una dirección transversal al haz debe compensarse con el momento total en el sentido opuesto. Si no lo hace, puede suponerse que el desequi-

**6. SEIS INTERACCIONES HIPOTETICAS podrían conducir al descubrimiento en el SSC de algunas nuevas partículas postuladas por la teoría. El gluino y el squark son los llamados compañeros supersimétricos del gluon y el quark. Si ambas partículas existieran y la masa del gluino fuera mayor que la masa del squark, las partículas se desintegrarían, según ilustran los dos diagramas de arriba. Si la masa del squark fuera la mayor, las partículas podrían desintegrarse como se muestra en los dos diagramas centrales. En ambos casos, la desintegración del gluino se ofrece a la izquierda. En los diagramas de abajo se dibujan dos tipos de desintegración del bosón de Higgs. El proceso inferior derecho predominaría si la masa del bosón de Higgs fuera de unos 50 GeV; el inferior izquierdo sería el típico si la masa fuera de unos 200 GeV o más.**

librio se produjo por la presencia de un neutrino, partícula cuya interacción con la materia resulta tan rara que casi nunca se detecta.

El primer disparador del SSC puede exigir sólo el paso de un electrón o muon moderadamente energético. Los niveles más elevados de toma de decisiones pueden recoger información cada vez más refinada de varias clases de detectores acerca de la energía y la trayectoria de las partículas emergentes. Para ciertos tipos de disparadores, bastará un detector que rodee sólo parte del lugar de la colisión. Sin embargo, en la mayoría de casos, se necesitan detectores polivalentes que envuelvan el punto de la colisión tanto cuanto sea posible. Podemos ilustrar mejor las funciones de un detector describiendo las principales características de un detector polivalente.

## Detectores polivalentes

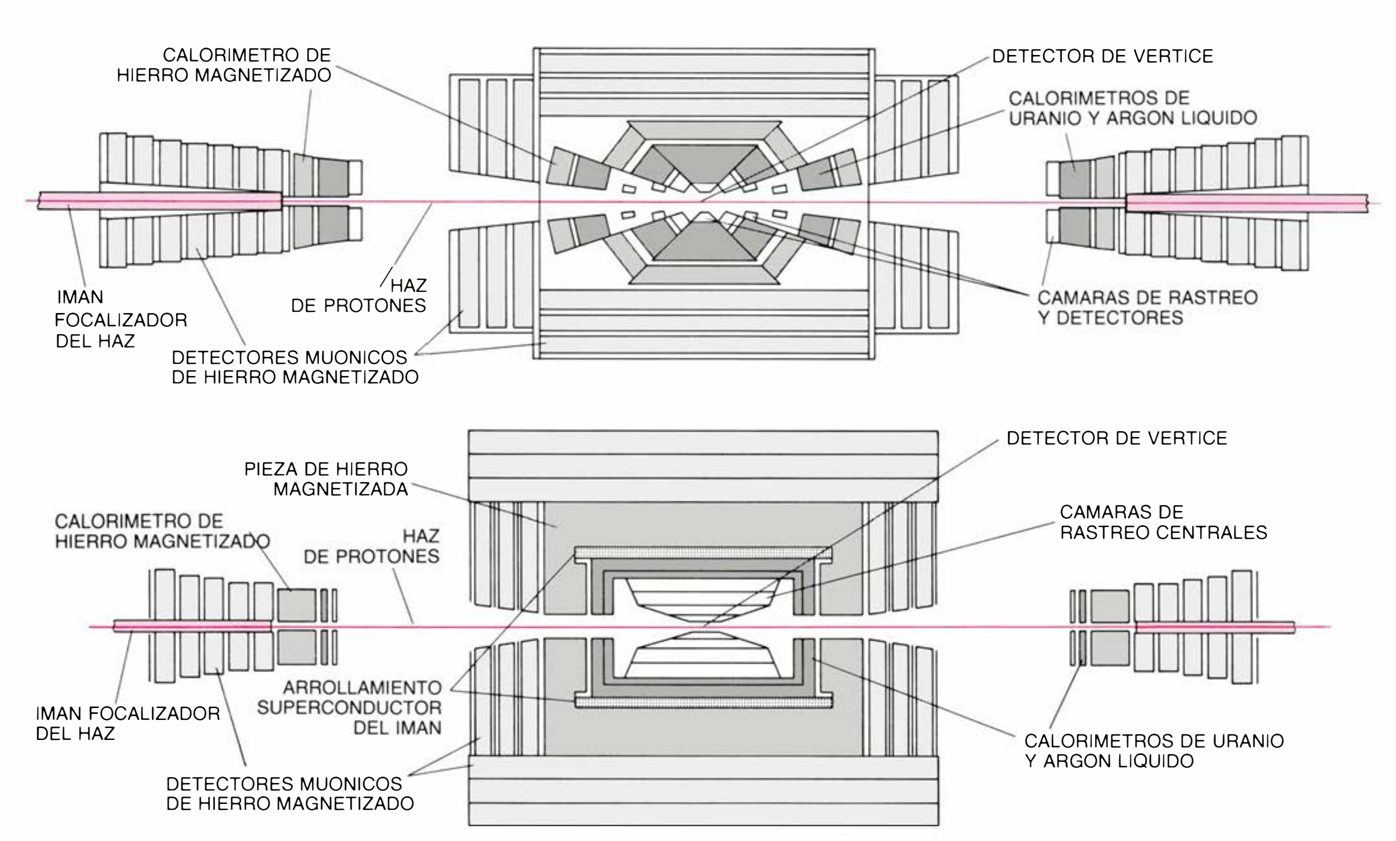
Para múltiples mediciones, el detector polivalente con un campo magnético en su región central ofrece evidentes ventajas, pero exige un precio a pagar. La gran complejidad del proceso de detección de las trayectorias y el material inerte necesario para los arrollamientos del imán dificultan las mediciones; los imanes encarecen notablemente el coste. Por consiguiente, un diseño óptimo de los detectores iniciales del SSC podría incluir un detector con un campo magnético y otro detector sin dicho campo.

Los componentes de que constan tales detectores son bastante similares [*véase la figura 7*]. El componente más interno será probablemente un detector de vértice, diseñado para medir las trayectorias de las partículas, con la máxima precisión posible, cerca del vértice, o punto de colisión de los protones. Las partículas avanzan entonces hacia afuera, pasando por una cámara central de rastreo que mide la dirección y, si hay un campo magnético, la curvatura de las trayectorias de las partículas cargadas. Fuera mismo de la cámara de rastreo, o quizá solapados con ella, se alojarán detectores que identifiquen los electrones.

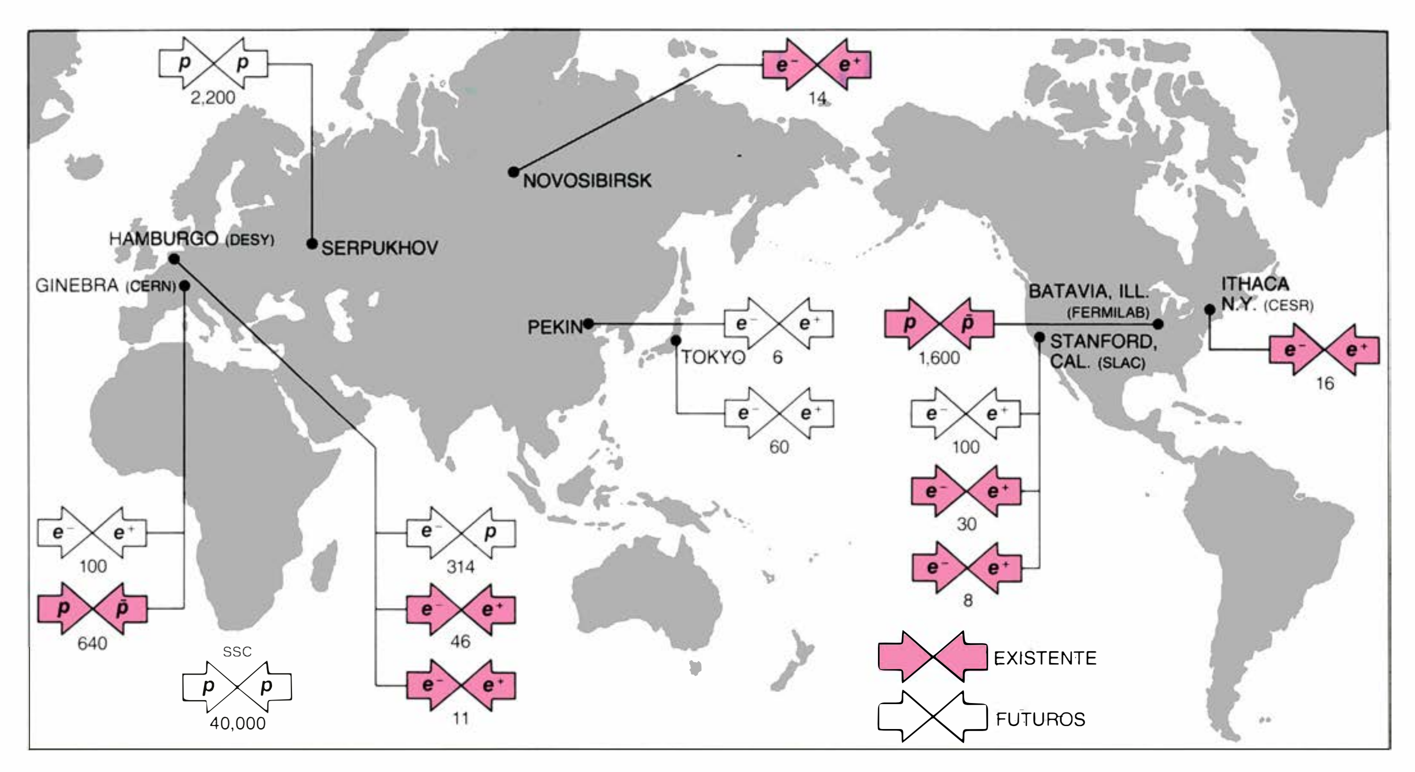
El grupo siguiente de detectores activos son los calorímetros. Miden la energía total de todas las partículas emitidas en direcciones mayores que cierto ángulo crítico con respecto a la dirección de los haces de protones incidentes. Gracias a los calorímetros se espera inferir la presencia de neutrinos de alta energía producidos en colisiones "duras", o frontales, entre los constituyentes de los protones. Recordemos que, en principio, los neutrinos pueden detectarse hallando una violación manifiesta de la conservación del momento transversal. Para las partículas que se mueven a velocidad próxima a la de la luz, el momento equivale a la energía. En la práctica, puede detectarse una ausencia de balance en el momento transversal de los productos de la colisión, detectando la ausencia de balance de su energía transversal.

La medida de la energía transversal puede hacerse midiendo la energía individual de cada partícula que sale de la colisión. Pero esa medición resultaría prohibitivamente difícil. Cada choque producido en el SSC puede generar más de 100 partículas, y se espera que muchas de éstas salgan en chorro o en forma de densos paquetes, lo que entorpece todavía más su distinción. Además, una técnica estándar para determinar la energía de una partícula cargada, basada en la medida de la curvatura de su trayectoria en un campo magnético, no puede aplicarse a una partícula neutra.

Los calorímetros miden la absorción




**7. DETECTORES POLIVALENTES DE PARTICULAS en el SSC. Tendrían por función cubrir el espacio en torno al lugar de colisión de los dos haces de protones. El detector superior carece de campo magnético alrededor de las cámaras centrales de rastreo; el inferior incorpora un imán superconductor que genera un campo a lo largo de la dirección del haz. Ambos detectores poseen una parte central y una parte hacia adelante; la parte hacia adelante está diseñada para detectar las partículas que se muevan casi paralelamente al haz. La parte interior de cada detector medirá diversas propiedades de cada traza cargada; tales propiedades pueden utilizarse después para reconstruir la trayectoria de una partícula. La parte exterior de cada detector es un calorímetro que mide la energía total depositada en cada segmento. El subsistema más externo de cada detector está formado por capas magnetizadas de hierro intercaladas con cámaras de rastreo; este subsistema identifica los muones y mide sus energías. El calorímetro absorbe las demás partículas, a excepción del neutrino.**




**8. ACELERADORES DE HACES COLISIONANTES dominan actualmente la tecnología de los experimentos de la física de altas energías. El mapa muestra tanto las máquinas existentes como las proyectadas. Los aceleradores se identifican de acuerdo con las partículas cuya colisión provocan: $e^-$ para el electrón, $e^+$ para el positrón, $p$ para el protón y $\bar{p}$ para el antiprotón. Los números indican la energía total máxima de colisión, medida en miles de millones de electronvolt.**

total de energía en algún medio, sin distinguir la contribución individual de cada partícula. A pesar de esta limitación, la dirección de los depósitos de energía puede determinarse segmentando los calorímetros. Así, la energía transversal total y la dirección de los chorros puede medirse bastante bien; a partir de esos datos puede determinarse si la colisión generó o no un neutrino.

El calorímetro de un detector polivalente está en el primer estrato del sistema de detección de muones, que a menudo son señal importante de sucesos de interés. En su exterior, habrá varias capas de hierro magnetizado, con cámaras de rastreo intercaladas entre ellas. Propósito del proyecto es permitir una redundancia en la medición del momento de los muones: desintegraciones de escaso interés pueden simular fácilmente el paso de un muon energético.

## Las fronteras de la física de altas energías

Hemos resaltado que la creación y detección en el SSC de partículas que tengan masas por encima de 1 TeV extenderá el conocimiento de los procesos elementales más allá del ámbito del modelo estándar. En particular, incidirá en el problema fundamental del origen de la masa y en el de la rotura de la simetría en la teoría electrodébil. Sin embargo, el SSC contribuirá también a la resolución de cuestiones pendientes. Verbigracia, parece ser que existen tres generaciones de quarks y leptones. ¿Hay más? ¿Por qué los quarks y los leptones tienen masas progresivamente mayores en las sucesivas generaciones? ¿Se mantienen los quarks siempre ligados en forma de hadrones o asistiremos, por fin, a la aparición de quarks libres? ¿Hay relación entre quarks y leptones? Si así es, ¿en qué consiste? ¿Por qué las interacciones débiles muestran carácter levógiro o dextrógiro? ¿Son realmente entidades elementales los quarks y leptones? ¿Están formados por otros constituyentes básicos? ¿Continuará la mecánica cuántica aplicándose a escalas cada vez menores? ¿Puede abordarse la gravedad y la interacción de color desde la mecánica cuántica y unificarse quizá con las otras fuerzas conocidas?

A lo largo de los últimos diez años hemos asistido a varios intentos de generalizar la unificación parcial conseguida en la teoría electrodébil en una gran unificación de las interacciones electromagnética, débil y de color. Más cercana en el tiempo está la teoría de supercuerdas, que ha extendido la supersimetría a un formalismo matemático que podría instar, algún día, una unificación aún mayor: la interpretación unificada de las cuatro interacciones fundamentales, gravedad incluida. En el éxito de esas teorías de gran unificación se apoya la esperanza de que la física de partículas tendrá algo que decir acerca de los primeros momentos de la historia del universo y de que la cosmología deberá pronunciarse sobre la física de partículas.

Sostienen los astrónomos que el universo empezó cataclísmicamente en la gran explosión. En el universo primordial, sometido a una temperatura fabulosa, la plena simetría de las leyes de la naturaleza debió ser manifiesta en los instantes subsiguientes a la gran explosión. El estudio de lo muy grande y de lo muy pequeño convergen así en un punto de vista común: para continuar analizando la unidad y la simplicidad subyacentes de la naturaleza han de constuirse instrumentos que investiguen dominios energéticos progresivamente más altos. No podemos aventurar qué hallazgos nos aguardan. Pero la experiencia nos enseña que a menudo es el descubrimiento inesperado el que provoca una interpretación científica más profunda del mundo. El SSC, viable aunque ambicioso, nos transportará a dominios de energías desconocidas, donde los descubrimientos reales sólo pueden sospecharse, y nos dará acceso a los sucesos que tuvieron lugar casi inmediatamente después del principio del tiempo. La oportunidad y el reto que el supercolisionador superconductor nos ofrece excitarán a todos los que compartan el deseo de entender el mundo que nos da cobijo.

# Evolución vegetal simulada por ordenador

*Los ordenadores constituyen herramientas adecuadas para someter a ensayo las hipótesis, de carácter intrínsicamente estadístico, de la biología evolutiva. ¿Cómo recrean las tendencias fundamentales de la evolución vegetal?*

Karl J. Niklas

Se encuentran los biólogos evolutivos con un grave obstáculo a la hora de verificar sus teorías: a menudo, las hipótesis en que se basan no predicen situaciones concretas y de fácil falsación. Antes bien, tratan de describir tendencias estadísticas generales a las que se ajustarían poblaciones muy numerosas al cabo de intervalos temporales prolongados. Por otra parte, los organismos involucrados en esas hipótesis, y los ambientes que les rodearon, quizá hayan desaparecido hace centenares de millones de años, como es el caso de la evolución de las primeras plantas terrestres. Así, trazada ya en detalle la historia evolutiva de un linaje por paleobotánicos, paleozoólogos o genetistas, les resulta aún imposible ofrecer respuestas definitivas a ciertas cuestiones fundamentales: ¿por qué se dan los patrones de evolución observados? ¿qué proporción de lo que muestra el registro fósil debe atribuirse al azar y cuál a procesos bióticos claramente definidos y a presiones selectivas?

Disponemos hoy de una herramienta adecuada para verificar las hipótesis evolutivas: el ordenador. Capaz de manejar grandes conjuntos de datos y de efectuar cálculos e iteraciones con rapidez, el investigador puede elaborar con él modelos de procesos evolutivos complejos. Esas técnicas permiten someter a examen muchas de las nociones elaboradas intuitivamente por los biólogos acerca de las interacciones que se establecen entre organismos y ambientes. En ello han trabajado el autor y sus colegas Vincent Kerchner y Thomas D. O'Rourke, de la Universidad de Cornell, quienes han sometido a ensayo las consecuencias matemáticas de diversas nociones de la evolución vegetal.

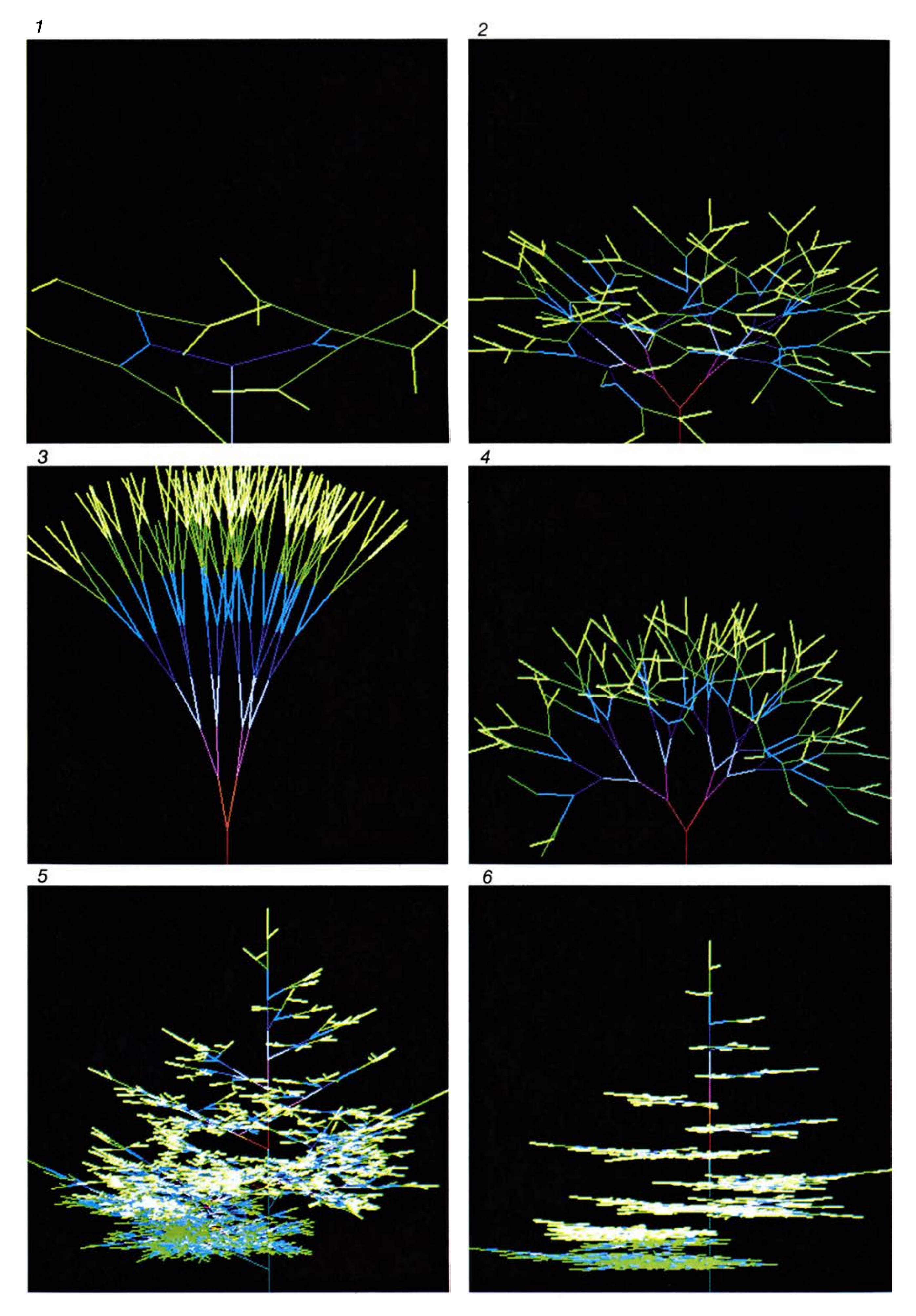
En las simulaciones por ordenador se examinan las hipótesis aplicando el método hipotético-deductivo. De acuerdo con el mismo, debe primero formularse una hipótesis. Seguidamente se deducen las diversas consecuencias a que dé lugar y se las compara con las pruebas recabadas por observación. Si las consecuencias concuerdan con las observaciones, la hipótesis recibe confirmación parcial. (Una hipótesis no puede "demostrarse" de esta forma, sólo aumentar la probabilidad de su validez.) Si alguna de las consecuencias deducidas de la hipótesis no concuerda con la observación, debe modificarse o rechazarse. El ordenador, que puede generar una gran "población de resultados" tras repetir muchas veces el mismo tipo general de cálculo, constituye una herramienta ideal para validar las hipótesis, intrínsicamente estadísticas, de la biología evolutiva.

Como se ha dicho, el primer paso es la formulación de hipótesis: enunciados sobre cuáles han sido los principales factores que han intervenido en la evolución vegetal. Con el ordenador se elabora entonces un modelo del comportamiento de plantas que presentan diversas características primitivas y se "evalúa" el éxito relativo de cada una de ellas en la solución de los problemas planteados por las presiones de selección supuestas en la hipótesis. Seguidamente, en esas plantas imaginarias se introducen cambios, pequeños o grandes, ("mutaciones") que quizá las hagan más eficaces, es decir, más capaces de hacer frente a las presiones asociadas a la hipótesis. Se evalúa a su vez la descendencia mutante y se somete a nuevas mutaciones, repitiéndose el ciclo una y otra vez. La etapa final consiste en comparar el patrón de evolución simulado con los patrones que muestra el registro fósil. De resultar satisfactoria la concordancia entre ambos, queda parcialmente confirmada la actuación de los factores cuya posible acción selectiva se supuso en la hipótesis.

Este método de simulación se apoya en dos supuestos principales, que cabe considerar los dos principios fundamentales de la biología evolutiva. El primero es que el carácter genético de los individuos de una especie, y por tanto el de la propia especie, cambia con el tiempo. El segundo es que existe cierto grado de continuidad entre un ancestro y sus descendientes. Los cambios registrados por una especie derivan, por consiguiente, de las presiones selectivas que actúan sobre gran número de pequeñas variaciones registradas entre los progenitores y sus descendencias.

Para el desarrollo de modelos de la evolución vegetal deben añadirse algunas hipótesis a los supuestos anterio-

**1. DESARROLLO EVOLUTIVO de las plantas terrestres primitivas, simulado por ordenador. La secuencia evolutiva obtenida se basa en la hipótesis de que la evolución respondía en parte a la necesidad de aligerar las tensiones mecánicas del patrón de ramificación, exponiendo a la vez la máxima superficie al sol. En la secuencia simulada, que exhibe algunas de las principales tendencias observadas en la evolución real de las plantas, el ejemplar más primitivo (*1*) es bajo y de escasa ramificación. (Los diversos colores indican las sucesivas generaciones de ramificación; las ramas de color verde brillante son las formadas más recientemente.) Las plantas más tardías tienden a poseer más ramas (*2*) y a crecer más en vertical (*3*); los patrones de ramificación verticales presentan más superficie fotosintética y pueden crecer por encima de la sombra proyectada por diversos obstáculos, como las plantas vecinas. La cantidad de sombra que se proyecta la planta a sí misma se reduce tendiendo grandes ángulos entre las ramas, lo que ensancha la planta (*4*). Los vegetales más avanzados (*5*) presentan un eje central simple a partir del cual se proyectan muchos ejes laterales. La capacidad de captación de luz de esta configuración se incrementa si los sistemas laterales de ramificación se aplanan (*6*). La similitud entre las tendencias producidas por tal simulación y las observadas en el registro fósil constituye una medida de la exactitud de las hipótesis en las que se basa aquella.**


1
2
3
4
5
6



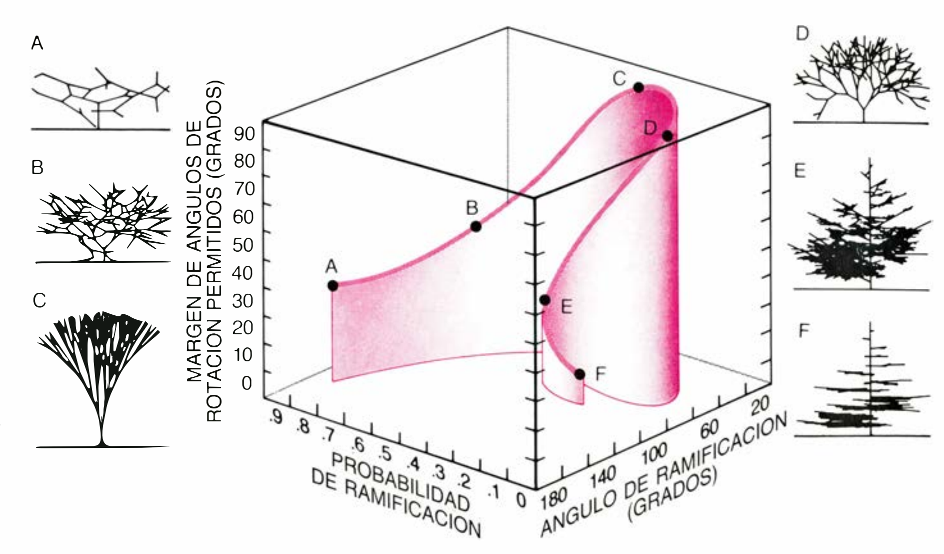

**2. SIMULACION DEL CRECIMIENTO de un vegetal; comprende una serie de estados a los que se llega por iteración. En cada ciclo el ordenador decide, basándose en la "probabilidad de ramificación" inherente de la planta, cuál de sus "ejes" se ramificará. Considera luego las demás características de la planta (por ejemplo los ángulos que forman las ramas) para determinar la dirección de los nuevos ejes, que se alargarán brevemente antes de subdividirse a su vez.**

res. Una es que la mayoría de las plantas pueden interpretarse cual soluciones estructurales de las limitaciones impuestas por el proceso bioquímico de la fotosíntesis. Así, podrá predecirse que las plantas con patrones de ramificación asociados a una máxima absorción de luz serán las de mayor éxito: los cambios de forma o de estructura interna que incrementen la capacidad de absorción lumínica conferirán ventaja competitiva.

Para competir con eficacia en pos de la luz y el espacio, los vegetales deben desempeñar también otras funciones; en particular, mantenerse erguidos, es decir, soportar las tensiones mecánicas que genera el crecimiento vertical. Una segunda hipótesis bien pudiera ser, por tanto, que la evolución vegetal respondió a la necesidad de reconciliar la capacidad de captar la luz con la de sostener estructuras verticales ramificadas. Una tercera hipótesis: la evolución de las plantas viene regida por su éxito reproductor, primándose los patrones de ramificación que permitan una mejor dispersión de semillas y esporas.

A partir de este conjunto de suposiciones e hipótesis, con un ordenador podrán simularse muchas de las tendencias principales observadas en las primeras fases de la evolución de las plantas terrestres. Las simulaciones efectuadas hasta ahora se refieren sólo a la primera fase de la proliferación de las plantas vasculares (plantas que contienen tejidos internos, xilema y floema, que transportan fluidos a la vez que colaboran en la sustentación vertical de la estructura). El período simulado no supera los 60 millones de años, desde hace alrededor de 410 millones de años hasta hace unos 350 millones de años. (La diversificación de las plantas con flores, el grupo de más reciente evolución, duró en cambio unos 100 millones de años.) No se han planteado hipótesis que consideraran la influencia de los herbívoros terrestres ni de los insectos polinizadores, pues su simulación resultaría mucho más compleja que las discutidas aquí. El empleo de ordenadores más potentes quizá permita algún día incorporar esos y otros factores que influyen en la evolución, así como los efectos derivados de cambios climáticos o de acontecimientos catastróficos.

Para simular la evolución vegetal deben desarrollarse técnicas matemáticas que cuantifiquen las ventajas competitivas ofrecidas por diversos caracteres. Una de tales técnicas determinaría cuanta luz puede interceptar una planta con determinada estructura.







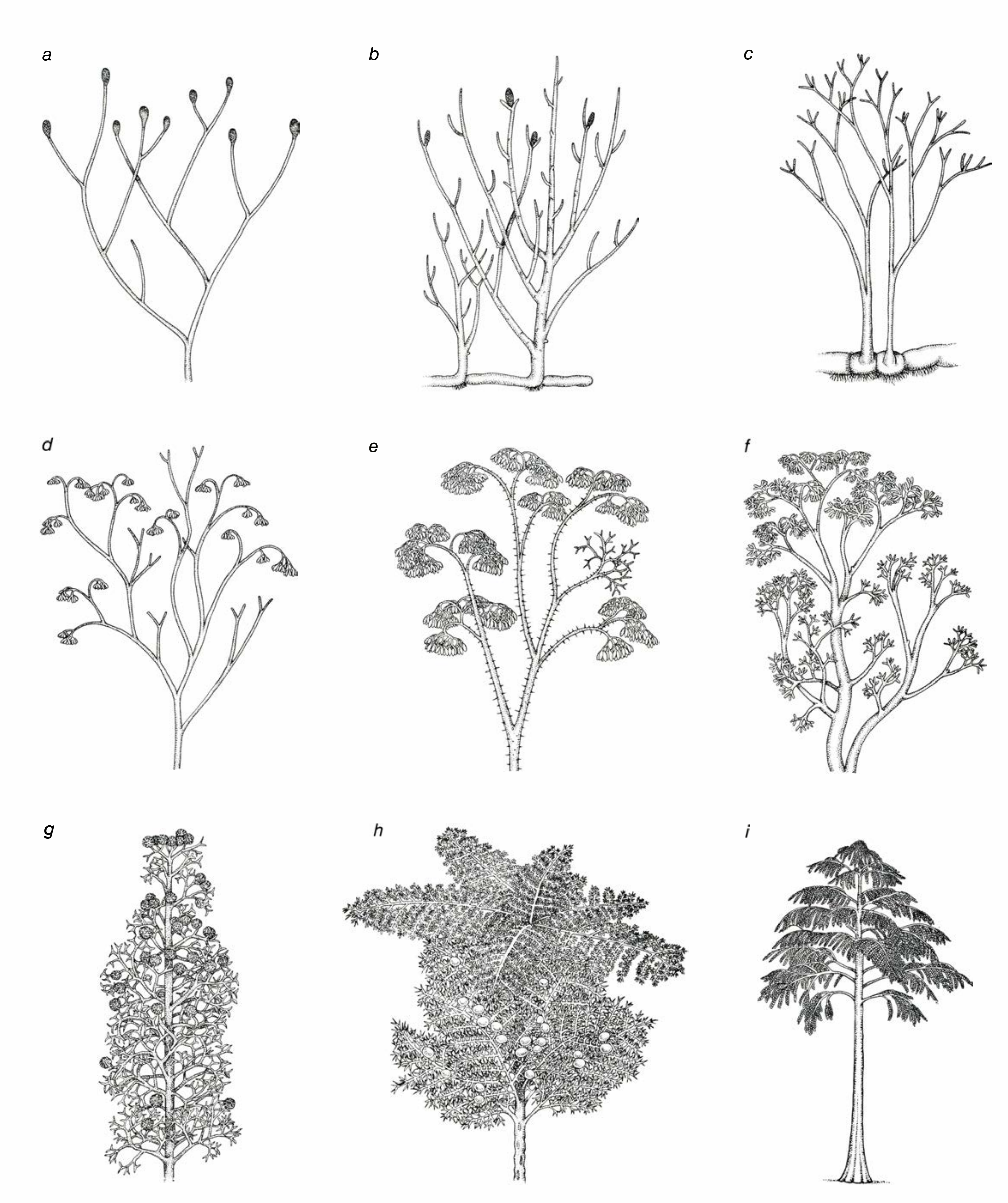
**3. ESPACIO CUBICO que engloba el "universo" de posibles patrones de ramificación. Cada una de las tres dimensiones mutuamente perpendiculares del cubo representa una característica morfológica particular y, cada punto del cubo, una "especie" vegetal dotada de una combinación única de características. Las tres cualidades aludidas son la probabilidad de ramificación, el ángulo de ramificación (el ángulo que tienden las ramas adyacentes) y el margen de ángulos de rotación tolerado (el ángulo formado por el plano de un par de nuevos ejes y el plano equivalente de la anterior generación de ramificación, definido por el eje a partir del cual han crecido los nuevos ejes y su hermano). El ordenador puede simular una "trayectoria evolutiva" (*color*) a partir de la planta que más se parezca a una especie primitiva. Se examinan las vecinas de la planta en el cubo y se determina cuál es la más eficaz de acuerdo con las hipótesis seleccionadas por el investigador. Esa planta vecina constituirá entonces el siguiente punto de la trayectoria, con el que el ordenador comparará todas sus vecinas en busca del ejemplar más eficaz. Se repite el procedimiento hasta que el programa localiza la planta dotada de más eficacia que cualquier otra.**

Por lo que sabemos, las primeras plantas vasculares, que dieron lugar al grueso de la actual flora de la tierra, carecían de hojas y presentaban ejes fotosintéticos verticales. (Los tallos, que son ejes portadores de hojas, evolucionaron posteriormente.) Los ejes tendían a crecer longitudinalmente, mediante la adición de nuevas células producidas por un grupo celular del ápice, o cono vegetativo; en general era escaso el número de nuevas células que se añadían a la sección, de ahí que el diámetro máximo tuviera cierto límite. La ramificación se producía al escindirse un eje en dos, de crecimiento independiente.

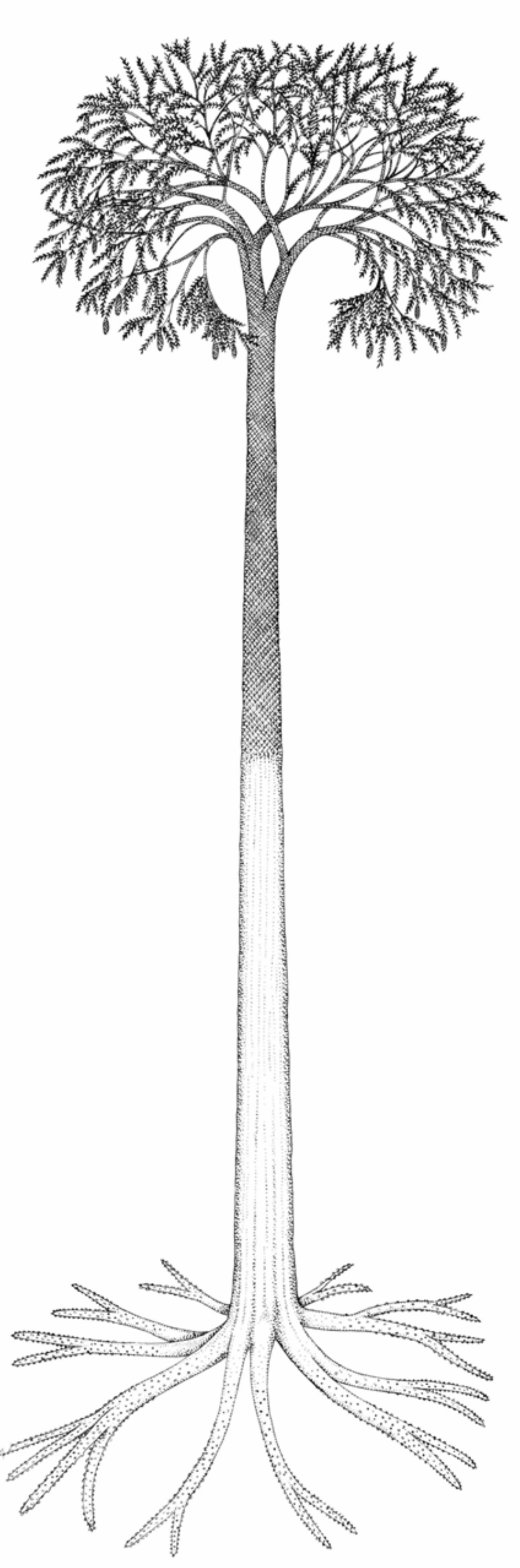
Desprovistas de hojas, los órganos fotosintéticos principales de muchas de esas plantas eran los ejes, relativamente rígidos. De ahí que tanto su geometría de ramificación como la forma en que las superficies de sus ejes se proyectaban hacia el sol constituyeran los factores más importantes a la hora de determinar su capacidad de captar luz. A diferencia de las hojas de las plantas actuales, los ejes de los primeros vegetales vasculares no podían agitarse con la brisa ni seguir el movimiento del sol a lo largo del día. Para evaluar con un ordenador la captación de luz solar por parte de las plantas primitivas bastará, por tanto, determinar la cantidad total de luz solar que se proyecta sobre una estructura tridimensional estática a medida que cambia la posición del sol.

Debe cuantificarse también la estabilidad mecánica de los ejes verticales. Los patrones de ramificación que crecen principalmente en dirección vertical captan mejor la luz, puesto que superan las sombras proyectadas por diversos obstáculos: rocas, montículos u otras plantas; pero ello obliga al vegetal dotado de esa estructura vertical a sustentar la carga mecánica concomitante. El principal soporte de la carga de los árboles actuales es el tronco leñoso, que ensancha continuamente su grosor a medida que se van agregando en su interior nuevas capas cilíndricas de células. Carentes de ese leño secundario, la capacidad de las plantas vasculares primitivas para soportar las tensiones mecánicas dependía sólo de su crecimiento primario: los ejes, relativamente delgados y de grosor casi constante, producidos por adición de nuevas células a los ápices de ejes preexistentes.

Resulta bastante fácil calcular la carga asociada a un patrón de ramificación, así como la capacidad del mismo para soportarla, cuando se conocen el peso, el tamaño y la orientación de los ejes. Los programas que generan los patrones de ramificación calculan también la tensión y el momento de torsión de un patrón de ramifica-




**4. EVOLUCION DE LAS PLANTAS PRIMITIVAS. Destaca la persistencia de ciertas tendencias (incremento del número de ramificaciones, emergencia de un eje principal y aplanamiento de los sistemas de ramificación) observadas también en trayectorias evolutivas simuladas. (Muchas de las plantas representadas no están directamente emparentadas: ilustran tendencias evolutivas generales, no la historia evolutiva de un linaje particular.) Una de las primeras plantas vasculares fue *Steganotheca striata* (*a*), planta escasamente ramificada que vivió hace unos 450 millones de años. Podrían contarse entre sus descendientes *Ryhnia gwynne-vaughanii* (*b*) y *Horneophyton lignieri* (*c*), ambas con estructuras de ramificación más extensas que las de *Steganotheca*. Las tres forman parte del grupo de las riniofitas. El siguiente grupo principal, que vivió aproximadamente hace 380 millones de años, es el de las trimerofitas, con *Trimerophyton robustius* (*d*), *Psilophyton charientos* (*e*), *Psylophiton dawsonii* (*f*) y *Pertica quadrifaria* (*g*). (Se desconocen las relaciones taxonómicas exactas y la secuencia de ocurrencia de esas trimerofitas.) Las trimerofitas parecen tener un tallo vertical principal; de hecho, ese tallo principal aparente constituye el resultado estadístico de gran número de "ramificaciones desiguales", en las cuales un miembro del par de nuevos ejes crece en una orientación más cercana a la del eje de procedencia del par. Un descendiente de las trimerofitas con aspecto de helecho, *Rhacophyton ceratangium* (*h*), posee, en efecto, un tallo vertical prin-**



**cipal; presenta también prefrondas planas, apéndices con aspecto de hojas que señalan un avance respecto de las ramas aplanadas. Otro posible descendiente de las trimerofitas es *Archaeopteris* (*i*). Los ejemplares de ese género poseían un tronco central portador de ramas laterales, que a su vez llevaban hojas. Un último linaje, el de los licopodios, parece haber evolucionado independientemente de riniofitas y trimerofitas. Los licopodios alcanzaron su cima en árboles licopodiales, como *Lepidodendron* (*j*).**

ción, igual que la cantidad de luz solar interceptada. Lo que permite cuantificar los factores implicados en el conflicto que se suscita entre la exposición de grandes áreas de tejidos fotosintéticos al sol y el soporte de las tensiones mecánicas resultantes.

Otro conflicto a considerar guarda relación con las sombras. Las plantas de mayor diámetro y más ramificadas están mejor dotadas para hacer sombra a las plantas vecinas y obtener así una ventaja sobre ellas. Sin embargo, esa ventaja ofrece también un aspecto negativo: probablemente se incremente la tendencia a que unas partes de la planta tapen a otras, lo que reduce la eficacia propia en la captación lumínica.

La hipótesis de que la capacidad de captación de luz constituyó una fuerza directriz de primer orden en la evolución temprana de las plantas vasculares es, ciertamente, una simplificación. En la flora actual se descubre una plétora de plantas que toleran la sombra, y hay razones para creer que entre los primeros vegetales vasculares también las había de ese tipo. Con todo, se advierten en el registro fósil tendencias a largo plazo que encajan con la hipótesis de que la competencia por la luz resultó primordial en la evolución de las plantas terrestres. Entre esas tendencias se cuentan el incremento de estatura en los grupos de plantas de evolución sucesiva, la transición a un patrón de crecimiento caracterizado por un solo eje vertical portador de ejes laterales y la aparición de sistemas de ramificación laterales aplanados, que en algunos linajes mimetizaron la función de las hojas. Se han recreado con cierta precisión esas tendencias mediante simulaciones en las cuales los patrones de ramificación "crecen" de acuerdo con las formas que definen ciertas características "genéticas". Durante cada fase del crecimiento simulado, la planta atraviesa distintas fases de alteración. Primero, cada eje se prolonga ligeramente. El ordenador selecciona entonces, a partir de unas "probabilidades de ramificación" predeterminadas genéticamente, cuántos ejes se ramificarán y cuáles. En los que se ramifican, las direcciones de los nuevos ejes están determinadas por dos factores "genéticos". Uno es el "ángulo de ramificación" entre los nuevos ejes. El otro es el ángulo de rotación: el ángulo que tienden el plano definido por los dos nuevos ejes y el plano formado por la generación de ramificación anterior en el eje de origen [*véase la figura 3*].

En las simulaciones más simples de patrones de ramificación, el ángulo de ramificación es el mismo en todos los puntos de divergencia; al igual que la probabilidad de ramificación, se trata de una de las características "genéticas" de los patrones. El ángulo de rotación, por otro lado, puede variar dentro de cierto margen, característico de la planta simulada. Los ejemplares a los que se dota de una gama amplia de ángulos de rotación tienden a ser redondos y arbustiformes; aquellos en los que el ángulo de rotación no se aparta mucho de 0 los grupos de ramas tienden a quedar en planos paralelos. Determinadas ya por el ordenador la localización y la dirección de los nuevos ejes, crecen todos un poco más y se repite el proceso de ramificación. Cumplidos 10 ciclos se considera que las plantas han satisfecho el crecimiento. El ejemplar al que, durante la simulación, se le ramifiquen todos los ejes en los 10 ciclos de crecimiento contendrá, en última instancia, 1024 elementos axiales. Podrían simularse plantas mayores, pero ello obligaría al uso de ordenadores con más memoria. Bastan 10 ciclos de ramificación para imitar el tamaño y la forma de la mayoría de las plantas fósiles que nos interesa.

Las tres características consideradas (probabilidad de ramificación, ángulo de ramificacion y ángulo de rotación) definen un universo de posibles patrones de ramificación que describe la morfología de las primeras plantas terrestres. La manera más simple de imaginar este universo es el cubo. Cada una de las tres dimensiones mutuamente perpendiculares del cubo corresponde a una de las tres características fundamentales, de manera que todo punto del cubo representa un grupo de vegetales que comparten un conjunto particular de atributos. (Dada la naturaleza aleatoria de la ramificación, aun poseyendo exactamente las mismas características matemáticas pueden diversos ejemplares diferir ligeramente en su morfología.)

Por ejemplo, supongamos que la dimensión vertical del cubo representa el ángulo de ramificación. Todo punto cercano al extremo superior del cubo correspondería a una "especie" cuyas ramas tienden entre sí un ángulo abierto, mientras que los situados cerca de la base del cubo representarían especies con ángulos de ramificación pequeños. En un vértice del cubo la probabilidad de ramificación es casi nula, los ángulos que forman las ramas son muy pequeños y, casi nulos también, los ángulos de rotación. Los puntos algo alejados de ese vértice en alguna

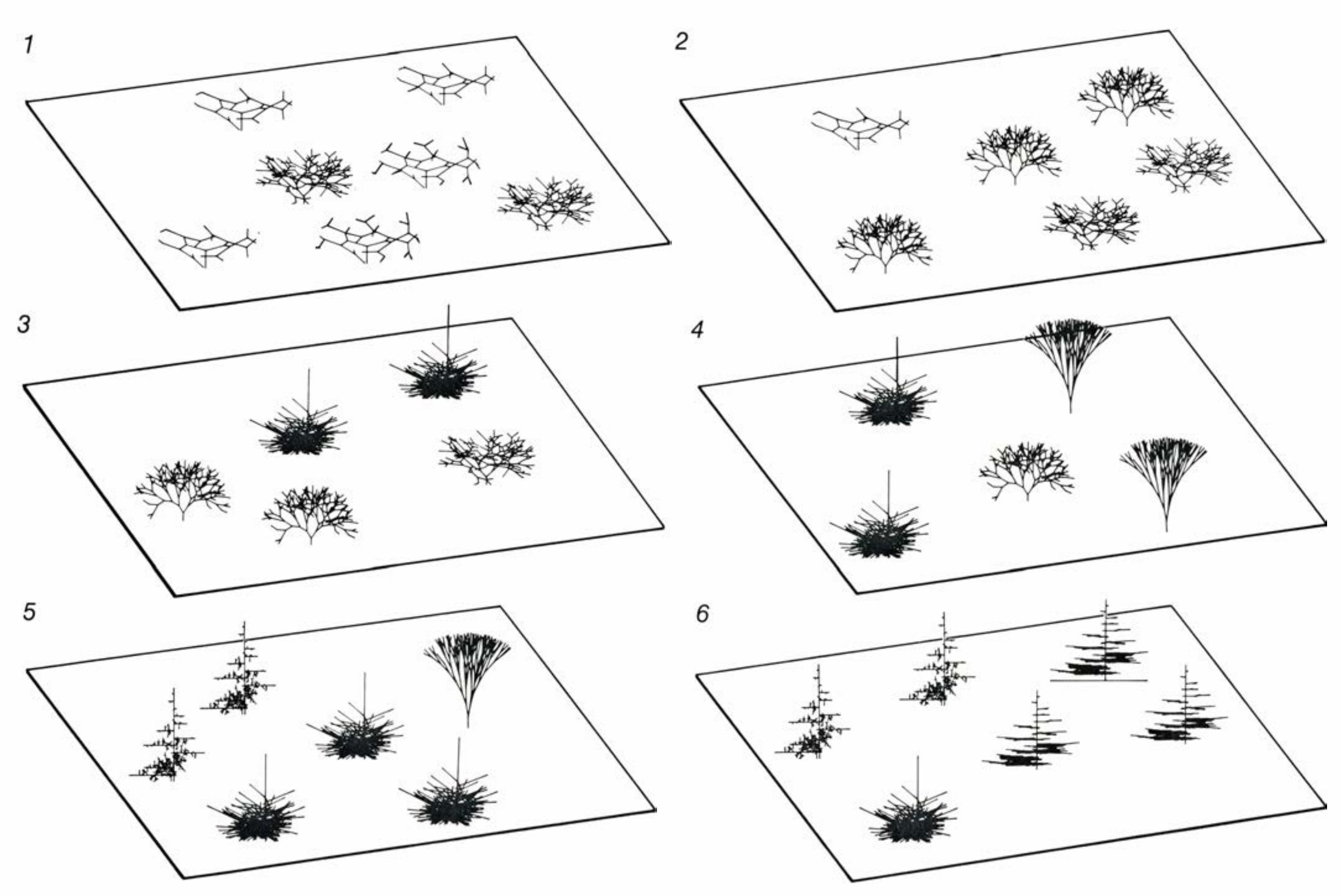
de las tres direcciones principales corresponden, respectivamente, a plantas con más ramas, con mayores ángulos de ramificación o con ramas que crecen en más direcciones. La combinación de las tres variables representada por los puntos del cubo entero incluye virtualmente todas las geometrías de ramificación de las plantas vasculares primitivas.

Por supuesto, el universo tridimensional constituye una simplificación excesiva de la geometría de las plantas. En realidad, además de los tres considerados aquí, muchos otros factores influyen en la forma de las plantas. Pueden simularse universos pluridimensionales de la morfología vegetal incorporando factores como la capacidad de cambiar la longitud de los ejes, o su grosor, así como la de producir ejes con distintas probabilidades de ramificación. En particular, una simulación más elaborada consideraría una variable que contemplara el fenómeno denominado ramificación desigual: en cada punto de ramificación uno de los nuevos ejes puede desviarse más que el otro de la orientación del eje a partir del cual crecen ambos. En ese caso existen dos ángulos de ramificación, cada uno de los cuales representa el ángulo tendido entre uno de los nuevos ejes y el eje original. Resulta difícil imaginar esos universos pluridimensionales, pero se simulan fácilmente con ayuda del ordenador. Por mor de simplicidad baste el universo cúbico, tridimensional, para describir las fases iniciales de la evolución vegetal.

En el cubo puede representarse cualquier "trayectoria evolutiva" simulada de las primeras plantas terrestres. En primer lugar se determina la localización del patrón de ramificación más primitivo (aquel que más se parezca al de las plantas más antiguas). Seguidamente, el ordenador selecciona, de entre los vecinos más cercanos a la planta en el cubo, la especie que resulte más eficaz en la captación de luz. (La eficacia de una especie se estima generando 10 ejemplares aleatorios y calculando la media de su rendimiento respectivo en la captación.) Se itera el proceso hasta que el ordenador identifica un conjunto de características que resulta más eficaz que cualquiera de los vecinos inmediatos de la especie en el cubo. La línea que una todos los puntos correspondientes a las especies vegetales seleccionadas constituirá una trayectoria evolutiva basada en la competencia por la luz.

Puede trazarse también una trayectoria que recoja la resultante del conflicto que se establece entre la eficacia en la captación de la luz y la tensión mecánica. En el cálculo de tal trayectoria por el ordenador, al examinarse los vecinos más cercanos de la planta para determinar qué especie ocupará el siguiente punto de la trayectoria se escoge la que muestra la razón más alta entre eficacia en la captación lumínica y momento de torsión, en vez de elegir directamente la más eficaz a la hora de captar luz.

Esas trayectorias evolutivas, tanto las que contemplan sólo la captación de luz como las diseñadas para reconciliar esa capacidad con el soporte de la tensión, presentan un notable parecido con las tendencias que arroja el registro fósil. El patrón de ramificación más primitivo se caracteriza por la presen-




**5. SIMULA EL JUEGO DE LA GUERRA** la evolución impulsada por la competencia. La ilustración muestra seis etapas del juego; en cada fase, una o varias plantas de las tres especies compiten por la luz y el espacio: se elimina la especie que capta menos luz y las dos restantes diseminan sus esporas. Muta la mitad de las esporas liberadas por la especie con más éxito, engendrándose una tercera, que habrá de competir en la siguiente etapa del juego. La capacidad de crecimiento depende de la cantidad de "luz" que recibe una planta (que a su vez depende de la cantidad de superficie que expone y del grado en que ella misma y otras plantas le hacen sombra). El número de esporas que disemina una planta es función del número de terminaciones de rama que posea; el área de dispersión de las esporas depende de su altura. El éxito en el juego conjuga la capacidad de captar luz, de hacer sombra a las plantas vecinas y de reproducirse.

cia de pocos niveles de ramificación, amplios ángulos entre los ejes y planos de ramificación casi verticales. El más avanzado posee grupos planos de ejes laterales unidos a un único eje vertical principal.

Las geometrías intermedias muestran, en la trayectoria evolutiva simulada, la transición de una geometría primitiva a otra más avanzada. En primer lugar se incrementa el grado de ramificación. Seguidamente, en cada ramificación sucesiva crece un eje más vertical que el resto, hasta convertirse en eje central. Finalmente, los ejes laterales resultantes (los que se proyectan del central) se sitúan en planos casi horizontales. Tal secuencia recoge los cambios orientados a maximizar la eficacia en la captación de luz y minimizar las tensiones mecánicas. Se da la coincidencia de que ese desarrollo se ajusta a la secuencia que maximiza la cantidad de sombra proyectada sobre las plantas vecinas y a la vez minimiza el grado de sombra proyectada sobre la propia planta.

Las tendencias evolutivas simuladas concuerdan razonablemente con las que se descubren en el registro fósil; ello no obstante, tal correspondencia entre tendencias simuladas y reales pudiera ser mera coincidencia sin relevancia biológica alguna. Quizás otras hipótesis den mejor cuenta del registro fósil que la noción de que las plantas responden a presiones en favor de la captación de luz y del soporte de tensiones mecánicas. No cabe duda, por ejemplo, de que el éxito de una especie depende en gran medida de su éxito reproductor. ¿Puede el ordenador examinar también esta hipótesis?

Cabe responder a esa pregunta simulando un "juego de la guerra" en el cual los patrones de ramificación con distintas capacidades de absorción de luz y potenciales reproductores compitan entre sí en un universo representado en la pantalla del ordenador. El objetivo del juego es declarar un "vencedor evolutivo" de acuerdo con las reglas del juego; éstas vienen a ser hipótesis sobre el supuesto proceder de la evolución.

A cada planta se le asigna un área de crecimiento, definida por su capacidad para proyectar sombra sobre sus vecinas. Se le asigna también un área de distinto tamaño, centrada en la primera, sobre la que puede diseminar unidades reproductoras, que en las plantas primitivas eran esporas de dispersión eólica. La mayoría de las plantas terrestres primitivas desprendían las esporas desde los extremos de sus ejes, razón por la cual el número de esporas que produce cada planta se considera proporcional al número de extremos de rama asociados a su patrón de ramificación; la talla determina el área de diseminación de las esporas. (Se ajusta ello a un criterio aerodinámico: la altura desde la que cae cualquier objeto ligero o pequeño se demuestra decisiva en la determinación de la distancia, desde su punto de origen, a la que la brisa lo transporta.) La altura de una planta no sólo depende del número de ramificaciones que posea, sino también de su ángulo de ramificación; las plantas que presentan menores ángulos de ramificación son las que alcanzan mayor altura. Ninguna de las esporas depositadas en un área sombreada, incluida la creada por la progenitora, germina. (Esta regla también es razonable: muchas de las plantas vasculares inferiores actuales pueden inhibir el crecimiento de su propia progenie.)

Las plantas con áreas de sombra solapadas se obstaculizan mutuamente en la captación de la luz. La tasa de ramificación, y por tanto de crecimiento, de cada planta viene determinada por la cantidad de luz solar que recibe. Lo cual depende a su vez, en parte, de su tendencia a evitar la proyección de sombras sobre sí misma, en parte de cuán intensamente le hagan sombra otras plantas y también de la cantidad de superficie fotosintética que presente al sol.

Al empezar la competición se esparce por el "campo de juego" el mismo número de esporas de los tres patrones de ramificación primitivos más diferenciados. Cuando todas las plantas han alcanzado 10 niveles de ramificación, el ordenador determina las especies cuyos miembros captaron menos luz solar y elimina todos sus ejemplares. Las dos especies restantes dispersan entonces sus esporas. Muta la mitad de las esporas de la especie cuyos miembros captaron más luz: se alteran ligeramente sus factores "genéticos" de manera que se incremente la capacidad de la planta para proyectar sombra, evitar el autosombreado, producir esporas para diseminarlas en un área grande. Crecen de nuevo los tres patrones resultantes (los dos ganadores en el "asalto" anterior y el nuevo patrón, mutante) y se repite la secuencia. Si queremos complicar el juego, habrá que introducir mutaciones no ventajosas, pero ello apenas modifica los resultados y, en cambio, alarga considerablemente la partida.

Incluso en un ordenador la partida puede alargarse fatigosamente; no ol-

videmos que hasta una breve escaramuza requiere miles de cálculos. El juego termina cuando se ha seleccionado la última mutación de entre el repertorio ofrecido por la trayectoria en estudio. Se declaran vencedoras las especies que sobreviven hasta el final. Con ello se trunca el juego artificialmente. La evolución real no sólo produce cambios del patrón de ramificación de las plantas, sino también de índole fisiológica, como los que afectan a la capacidad para tolerar la sombra o de producir leño secundario. La evolución real puede verse también afectada por acontecimientos externos, piénsese en las catástrofes naturales, que obligan a recomenzar el juego con nuevas proporciones de jugadores. Con todo, tan sencillo juego de la guerra genera tendencias muy similares a las halladas en el registro fósil, por lo que probablemente las reglas invocadas en él guarden relación con la biología de las plantas reales.

La trayectoria que destaca la capacidad de captación de luz y la que reconcilia el conflicto entre captar la luz y soportar las tensiones muestran pareja idoneidad en la selección de mutaciones que generan plantas con éxito en el juego de la guerra. En ambos esquemas, la geometría más primitiva presenta escasos niveles de ramificación. De hallarse los esporangios (las estructuras portadoras de esporas) en el ápice de los ejes, resultaría pobre la capacidad reproductora de ese tipo vegetal. En ambos esquemas, la geometría más avanzada consta de un solo eje, grande, portador de muchos niveles de ejes laterales. Los esporangios de tal planta, en caso de formarse en los extremos de los ejes, serían numerosos; por añadidura, se encontrarían a considerable altura del suelo, asegurándose la buena dispersión.

Las observaciones de ejemplares vivos indican que el tipo de planta que tiene éxito en el juego de la guerra lo alcanza también en la competencia natural. Los estudios de dinámica de poblaciones realizados en monocultivos señalan que los individuos grandes aventajan a los pequeños. A medida que crece la densidad de la población, se incrementa correspondientemente la mortalidad, pero muere una proporción mayor de plantas pequeñas que de plantas grandes. Al igual que en el juego de la guerra, se han observado esos fenómenos en muchas especies vegetales.

Las mismas presiones que generan esas tendencias en los monocultivos reales bien pueden haber actuado sobre la evolución temprana de las plantas vasculares. En las primeras fases de la invasión de la tierra firme por parte de las plantas probablemente se encontrara casi despoblado el nuevo hábitat; las densidades de ocupación serían bajas. Durante cierto tiempo se incrementarían localmente el número de individuos y de especies. Las que alcanzaran más altura obtendrían ventaja respecto de las demás. A medida que aumentó la densidad de las comunidades vegetales, los individuos, e incluso las especies, morirían de forma selectiva, favoreciéndose la supervivencia de las especies mayores.

En este contexto valdrá la pena destacar un punto aparentemente contradictorio. Debido a la proyección de sombras del propio vegetal sobre sí mismo, cualquier incremento de la talla reduce la eficacia de la captación de luz: la planta absorbe más luz, pero capta menos luz por eje. Sin embargo, tal pérdida de eficacia la compensa la tendencia de la planta grande a proyectar sombra sobre sus plantas vecinas y las esporas de éstas, de modo que la inclinación hacia el desarrollo de ejemplares mayores debió conferir ventaja a la especie, aun a costa de perjudicar la eficacia individual de sus miembros. Aparentemente, entonces, la mera presencia de muchos individuos de la misma especie puede favorecer tendencias evolutivas que ayuden a la especie a competir con otras.

Las simulaciones por ordenador han demostrado que las diversas hipótesis propuestas acerca de la evolución vegetal provocan efectos verificables en el registro fósil. No puede afirmarse que se haya demostrado la validez de las hipótesis, pero sí que se han confirmado parcialmente. Es más, las consecuencias, simuladas, de las diversas hipótesis resultan casi idénticas. Lo que hace imposible determinar la importancia relativa de los diversos criterios aplicados en su elaboración; sólo puede efectuarse esa valoración cuando dos hipótesis conducen a conclusiones contradictorias. Nos aguarda ahora la simulación de requisitos estructurales tan importantes como la capacidad de transportar fluidos y de disipar el calor. El biólogo evolutivo contempla con esperanza el futuro de las técnicas generales de simulación, aun cuando el empleo fructífero de cualquier herramienta requiere práctica y criterio y exige el desarrollo y refinamiento de la propia herramienta de trabajo.

# Imágenes mentales y sistema visual

*¿Qué relación guardan la imaginería mental y la percepción visual? En recientes investigaciones el autor ha comprobado que ambas comparten muchos de los procesos nerviosos del sistema visual*

Ronald A. Finke

Abundan quienes dicen ser capaces de formar imágenes mentales de un objeto que semejan el aspecto real del mismo. El acto de construir tales imágenes produce a menudo sensaciones visuales que parecen enteramente realistas. Figurémonos, por ejemplo, que estamos viendo un elefante. ¿Tiene una trompa curvilínea? ¿De qué color son sus colmillos? ¿De qué tamaño sus ojos? La mayoría de la gente sostiene que, para responder a esas preguntas, "revisa" una imagen mental de manera casi idéntica a si inspeccionara un elefante real.

Esas informales observaciones sobre la imaginería mental le llevan a uno naturalmente a preguntarse en qué medida estarán relacionadas entre sí dicha imaginería y la percepción visual. Sugieren, en concreto, que en la representación mental de imágenes acaso participen muchos de los procesos nerviosos internos del tipo de los que subyacen a la percepción visual, posibilidad ésta que tendría importantes implicaciones teóricas y prácticas. Si se lograra probar, pongamos por caso, que en el sistema visual humano la imaginería mental participa de mecanismos nerviosos comunes a la percepción visual, podrían empezar a establecerse con precisión las interacciones recíprocas de las imágenes y la percepción visuales. Lo cual permitiría explorar los diversos modos de acción de la imaginería para facilitar, intensificar o inclusive sustituir a la percepción visual.

Durante los últimos diez años mis colegas y yo hemos venido desarrollando diversas técnicas para investigar la relación funcional que se establece entre las imágenes mentales y la percepción visual. Dado que a veces los sujetos experimentales conjeturan qué es lo que debería ocurrir en un experimento sobre imágenes, hemos procurado que nuestras técnicas resulten lo bastante precisas para revelar sutiles correspondencias entre la imaginería y la percepción. Nuestro trabajo ha puesto de manifiesto que las imágenes mentales despliegan un abanico de propiedades visuales mucho más rico de lo que se esperaba, pero también que, en ciertos aspectos, la imaginería difiere de la percepción.

La introspección permite reconocer que las características de una imagen mental formada a partir de algo pequeño o que se halle a bastante distancia son más difíciles de distinguir que las de la que se forma a partir de algo grande o que esté a poca distancia. Trátese, por ejemplo, de imaginar una hormiga sobre un periódico y a un metro de distancia, y luego esa misma hormiga puesta en la punta de un mondadientes y a sólo un palmo de nuestros ojos. Sin duda alguna, se podrán "ver" mentalmente muchos más detalles de la hormiga (así su cabeza y los segmentos de su cuerpo) al imaginarla más cerca.

Stephen M. Kosslyn, de la Universidad de Harvard, estudió tal relación entre el tamaño de la imagen y el detallamiento de sus características valiéndose de sencillas técnicas de medida de los tiempos de reacción. Averiguó que los rasgos característicos de un animal imaginado, tales como los ojos y las orejas de un gato, podían detectarse más rápidamente cuando a los sujetos se les había pedido que se figuraran unas imágenes relativamente grandes o que las supusieran relativamente cercanas. Los experimentos se inspiraron en la común observación de que las características de objetos físicos reales pueden detectarse más deprisa cuando a esos objetos se los mira desde menos distancia.

Más recientemente, Howard S. Kurtzman y el autor realizaron experimentos en la Universidad de Cornell para medir con precisión hasta qué punto se puede captar, o distinguir, por separado las características de los objetos en la imaginería mental y en la percepción. Lo que en particular nos interesaba era saber cómo afectarían el tamaño de los rasgos característicos, su espaciación y su posición en el campo visual a la resolución o capacidad de distinguir unos detalles de otros. Predijimos que, a través de todas esas variaciones, la resolución visual de las imágenes mentales igualaría a la resolución de la percepción.

En la percepción visual la resolución disminuye progresivamente conforme se observa el objeto desde sitios cada vez más alejados del punto de fijación de la mirada. La cantidad de detalles distinguibles no resulta, empero, la misma en todas direcciones. Por regla general, la resolución disminuye más despacio a lo largo del eje horizontal del campo visual que a lo largo del eje vertical, y también más despacio por debajo del punto de fijación que por encima de él. Se sabe asimismo que cuesta más distinguir las barras entrecruzadas a medida que el enrejillado se hace más fino o, dicho con más precisión, según va en aumento su frecuencia espacial básica.

Nuestro procedimiento para medir

**1. PAISAJE PINTADO A LA ACUARELA por una escocesa ciega que trabaja a partir de sus imágenes mentales. La artista, Carolyn James, padece una modalidad muy grave de la enfermedad ocular conocida por retinitis pigmentosa. Tiene 42 años y se quedó ciega a los 21. Para pintar alinea ante sí, en un orden memorizado, 24 tarros con colores de acuarela. Palpando las diversas secciones del papel se informa, por el grado de humedad, de cuáles son las que acaba de pintar. Sus acuarelas suelen tener seis capas de pintura.**

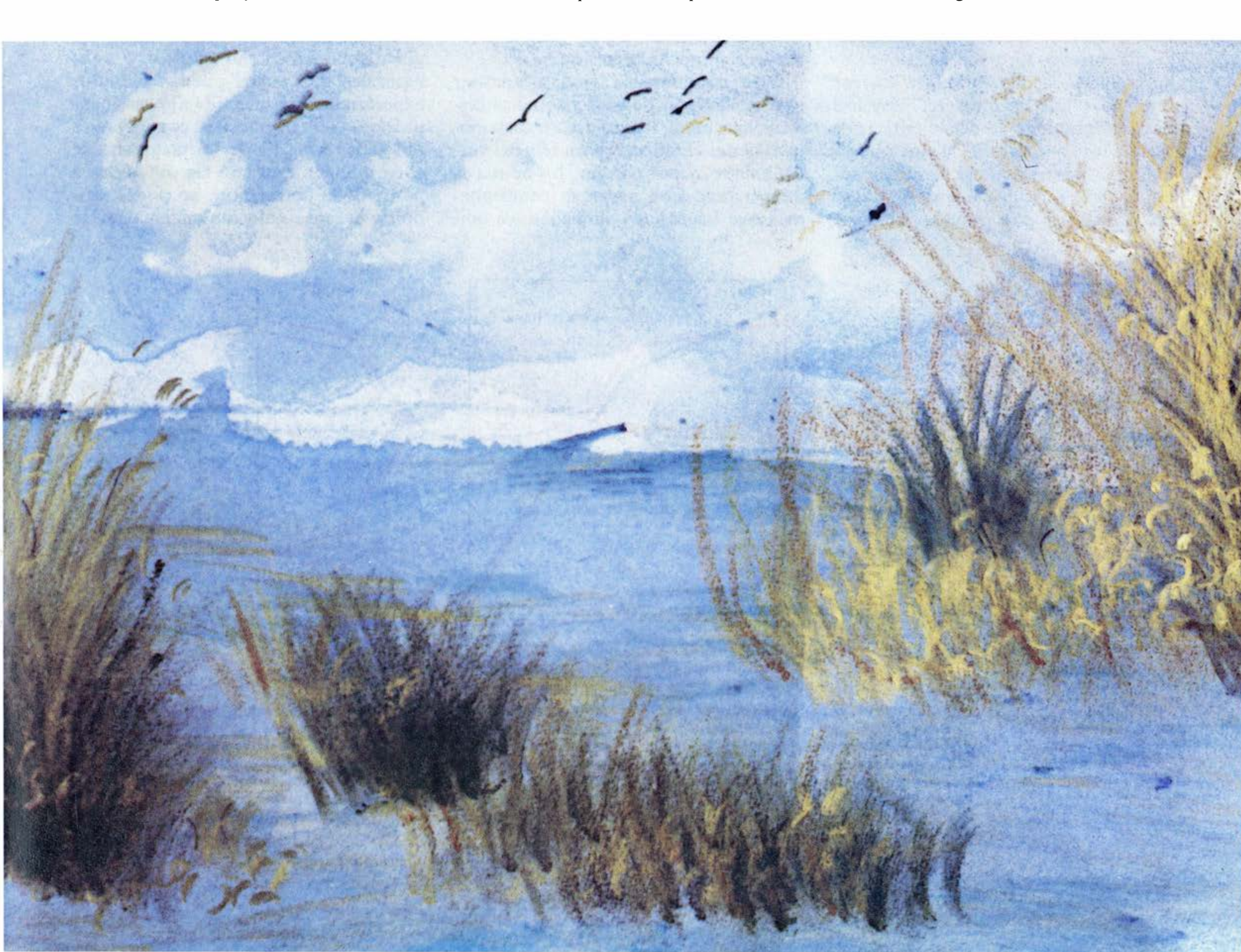
los límites de resolución de la imaginería mental se basaba en ciertas técnicas muy conocidas por los psicofísicos de la visión. Empezábamos mostrando a nuestros sujetos un disco plano en cuya mitad superior se había trazado una serie de barras verticales y, en la inferior, otra serie de barras horizontales. En ambos trazados era el mismo el grosor de las barras. Acto seguido les pedíamos que se formaran una imagen mental del disco y la proyectaran sobre el centro de una pantalla puesta directamente ante ellos y a partir de cuyo centro se expandían radialmente ocho líneas. Los sujetos indicaban hasta dónde les era posible apartar la mirada de sus imágenes al ir siguiendo cada una de las líneas sobre la pantalla, antes de que se les hiciese ya imposible distinguir las dos mitades del modelo imaginado. Declaraban que las retículas aparecían borrosas e indistintas a medida que se las imaginaba más cerca de la zona visual periférica, y que luego, rebasado cierto punto, no se las podía ya distinguir. Por comparación se obtuvieron los mismos juicios en unas condiciones perceptuales en las que el mismo disco se proyectaba realmente sobre el centro de la pantalla y frente a los sujetos.

Repetimos el experimento con tres discos diferentes. Las barras trazadas en el segundo eran tres veces más finas que las del primero y, las del tercero, tres veces más finas que las del segundo. Por término medio, las amplitudes de la resolución visual disminuyeron al aumentar la frecuencia espacial (con la mengua del grosor de las barras), y resultaron virtualmente idénticas tanto para las retículas imaginadas como para las observadas. Las formas de los campos imaginario y perceptual eran también muy similares: la resolución decrecía más despacio a lo largo del eje horizontal que a lo largo del vertical, así como más lentamente por encima del punto de fijación que por debajo de él.

Realizamos después un experimento de control en el que mostramos a otro grupo de sujetos el conjunto original de los tres discos y les pedimos que predijeran los campos de resolución de las imágenes. Pues bien, sus predicciones discreparon considerablemente de lo que teníamos observado, cosa que desautorizaba cualquier fácil explicación "conjetural" de nuestros primeros hallazgos.

Interpretamos los resultados como prueba de que la discriminación de muestras presenta poco más o menos las mismas limitaciones en la imaginería mental que en la percepción visual. Proponemos, además, que esas comunes limitaciones probablemente se impongan en los niveles de procesamiento de muestras del sistema visual, donde las propiedades de ciertos mecanismos nerviosos quizá limiten la capacidad para resolver rasgos visuales pequeños o poco espaciados. Kurtzman y yo hemos desarrollado otros experimentos que concuerdan con tal resultado. No hemos hallado correspondencia entre juicios de imágenes y juicios de objetos que incluyan diferentes cantidades de contraste visual, o relativa brillantez, entre los rasgos. Considérase que esos aspectos de la percepción se rigen por tipos de procesos nerviosos más primitivos, que operan en niveles inferiores a aquellos en los que tiene lugar el tratamiento de las muestras.

Nuestras averiguaciones sobre la resolución de las imágenes se basaron en



imágenes mentales de muestras planas, de dos dimensiones. Pero la imaginería mental típica es tridimensional: representa los objetos en relieve, tal como se los ve desde distintas perspectivas. Cuando la mayoría de la gente se imagina su cuarto de estar, por ejemplo, puede "ver" mentalmente que determinadas piezas del mobiliario se hallan en frente de otras, según en qué parte de la habitación se imagine hallarse.

Para investigar las propiedades tridimensionales de las imágenes, Steven Pinker y el autor propusieron a los sujetos que se formaran imágenes de una configuración de objetos situados en el espacio y las rotaran mentalmente. En la realidad, al contemplar esas configuraciones tridimensionales desde diversas perspectivas los objetos parecen cambiar de posición relativa en profundidad a medida que se mueve el observador, o cuando se rota la configuración.

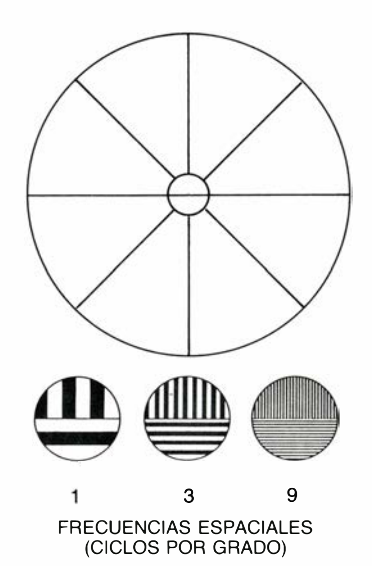
Recordemos, por ejemplo, el caso del tiovivo. Probablemente hayamos advertido que los que van montados en él parecen cambiar de posición con respecto a la nuestra según gira el tiovivo, formándose a veces figuras bidimensionales que nos resultan familiares, de manera análoga a como una constelación de estrellas forma en ocasiones figuras planas reconocibles. En los experimentos que hicimos en Harvard nos interesamos especialmente por averiguar si parecerían producirse semejantes tipos de figuras cuando los sujetos se imaginaran estar observando una configuración de objetos que diese vueltas. Los resultados arrojaron sorprendentes similaridades.

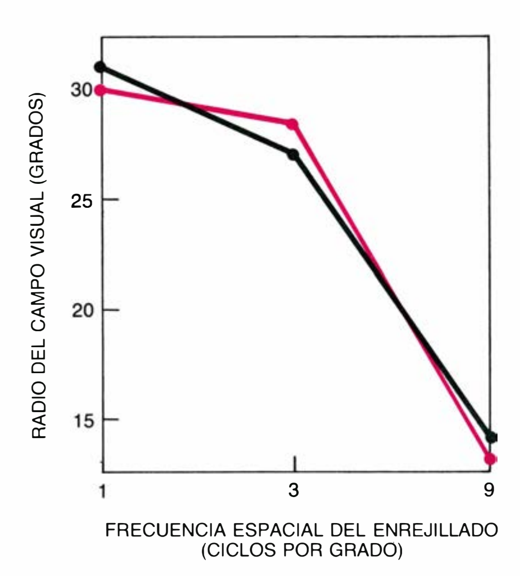
Pedimos a nuestros sujetos, primero, que se fijaran bien en las posiciones de cuatro animalillos de plástico colgados a distintas alturas dentro de un cilindro transparente y, después, tras haberlos retirado, que formaran imágenes mentales de cada uno de ellos. A continuación rotamos 90 grados el cilindro vacío y mandamos a los sujetos que se representaran imaginariamente el conjunto figural tal como aparecería desde el nuevo punto de vista. De haberse imaginado a los animales y haberlos hecho girar con toda exactitud según nosotros hacíamos rotar el cilindro, habría parecido que los animales formaban un paralelogramo.

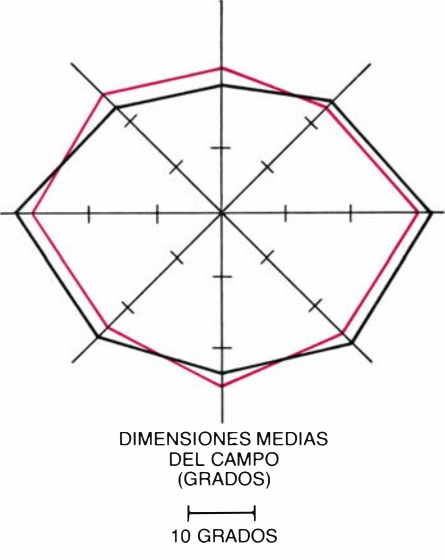
En todos los casos las representaciones de los sujetos revelaron que sus imágenes mentales habían descrito, tras la rotación, una figura muy parecida al paralelogramo, por más que la apariencia de los objetos vistos desde la posición original no sugiriese que fuera a emerger precisamente esa forma geométrica. Curiosamente, en las representaciones se produjeron pequeñas distorsiones, aunque sistemáticas. La naturaleza de éstas sugería que las rotaciones mentales no se habían efectuado con mucha precisión.

Otro experimento vino a apuntar una explicación de esas pequeñas distorsiones; en él, los sujetos giraban manualmente el cilindro para alinear verticalmente, y por parejas, los animales imaginados. Con sorpresa comprobamos que los sujetos alineaban los animales imaginados efectuando siempre giros más cortos de lo que era necesario cuando sí se encontraban presentes los animales. En otras palabras, los sujetos habían hecho rotar mentalmente sus imágenes anticipándose a su rotación manual del cilindro. Esta tendencia a hacer que la imagen se adelantase un poco explicaba las pequeñas distorsiones advertidas en las representaciones de las figuras emergentes efectuadas por nuestros sujetos. El experimento reforzaba, así, nuestra opinión de que podemos imaginar con exactitud las perspectivas visuales ofrecidas por muestras tridimensionales. Más aún, nos capacitaba para medir propiedades de imágenes mentales que de ordinario se les pasaría por alto a los sujetos no avisados.

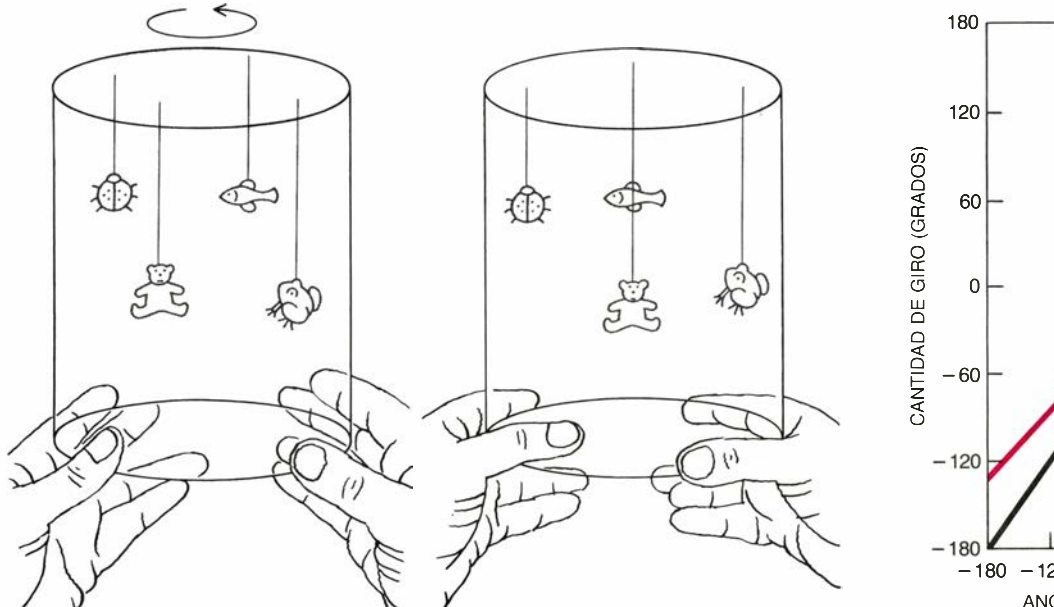
Sin embargo, haciendo patente que los sujetos no pueden adivinar el resultado de un experimento sobre imaginería mental no se excluye la posibilidad deque su realización se base en algún saber inconsciente acerca de los cambios del aspecto visual de los objetos, conocimiento éste que podría influir indirectamente en los juicios sobre imágenes. Cabe obviar ese problema instando a los sujetos a que imaginen fenómenos tan atípicos o antinaturales que nunca puedan haberlos experimentado antes. Si bajo esas condiciones las respuestas conductuales obtenidas a partir de la imaginería se corresponden aún con las obtenidas a partir de la percepción, no podrá atribuirse la representación imaginaria a la

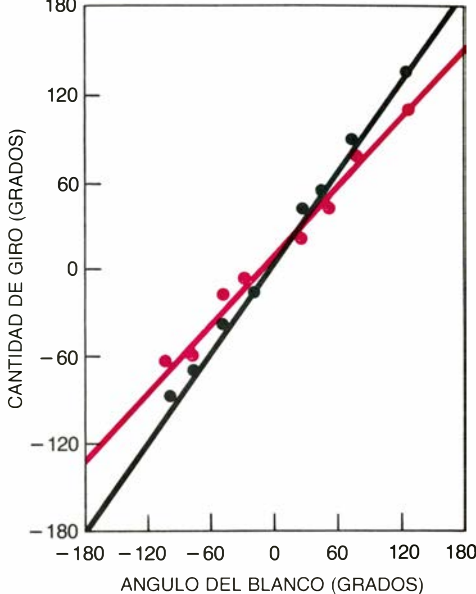



**2. RESTRICCIONES impuestas a la resolución visual, o capacidad de distinguir detalles; se midieron de forma experimental por medio de las tres figuras mostradas a la izquierda, abajo. A los sujetos sometidos al ensayo realizado por el autor se les mandaba formar imágenes mentales de cada figura y proyectar esas imágenes sobre el centro de un gran círculo situado ante ellos (*arriba, a la izquierda*). Los sujetos indicaban entonces hasta dónde podían apartar la vista de sus imágenes siguiendo cada una de las ocho líneas marcadas en el gran círculo sin dejar de distinguir las dos mitades de las figuras. Los campos visuales dentro de los cuales podían distinguirse los trazos de barras de las figuras imaginadas disminuían de tamaño al aumentar la frecuencia espacial o decrecer el grosor de las barras (*centro, línea de color*). Por lo general, esos campos se alargaban horizontalmente y eran mayores por debajo de la dirección de la mirada que por encima de ella (*derecha, figura de color*). Resultados idénticos (*centro, línea negra, y derecha, figura negra*) se obtuvieron cuando se proyectaron realmente las muestras sobre el gran círculo; lo que indica que en la imaginería y en la percepción se imponen similares limitaciones a la resolución de características.**






**3. PROPIEDADES TRIDIMENSIONALES de las imágenes mentales, exploradas por medio de un cilindro transparente. A los sujetos se les decía que se fijaran bien en las localizaciones de cuatro animalillos de plástico colgados a diferentes alturas dentro del cilindro, y luego, una vez retirados los animales, que formaran imágenes mentales de ellos. Tras girar 90 grados el cilindro vacío, los sujetos hicieron rotar también sus imágenes mentales (*izquierda*). Aunque la apariencia de los objetos en la posición de la visión primera no sugería que fuese a emerger un paralelogramo, con un animal en cada una de sus esquinas, los sujetos detectaron tal forma. Sin embargo, sus dibujos de esa forma geométrica mostraban pequeñas distorsiones, aunque sistemáticas. Otro experimento sugirió una explicación de la ocurrencia de esas distorsiones; en él, los sujetos giraban manualmente el cilindro para alinear verticalmente a pares los animales imaginarios (*centro*). Los sujetos rotaban siempre el cilindro un poco menos (*línea de color, a la derecha*) de lo que lo hicieron en una prueba similar en la que sí había animales (*línea negra, a la derecha*). En otras palabras, los sujetos habían rotado mentalmente sus imágenes antes incluso de girar a mano el cilindro.**

influencia de anteriores experiencias perceptuales.

En una serie de experimentos que llevé a cabo en el Instituto de Tecnología de Massachusetts, traté de proporcionar este tipo de evidencia al valor funcional de la imaginería considerando el posible papel que acaso desempeñaran las imágenes en la adaptación al prisma. Los prismas ópticos desplazan la ubicación aparente de los objetos. Numerosas investigaciones han demostrado que la gente se adapta pronto a observar el mundo a través de tales prismas, con tal que se la deje mover libremente y que pueda advertir sus errores. Cuando, una vez adaptados ya a los prismas, se les quitan éstos, cometen errores de movimiento en sentido opuesto, reflejándose los cambios habidos en su coordinación visual-motora. Mis experimentos demostraron que la adaptación al prisma puede producirse aun cuando la gente se mantenga quieta mirando a un punto fijo y se imagine cometer errores de movimiento como los inducidos típicamente por los prismas de desplazamiento.

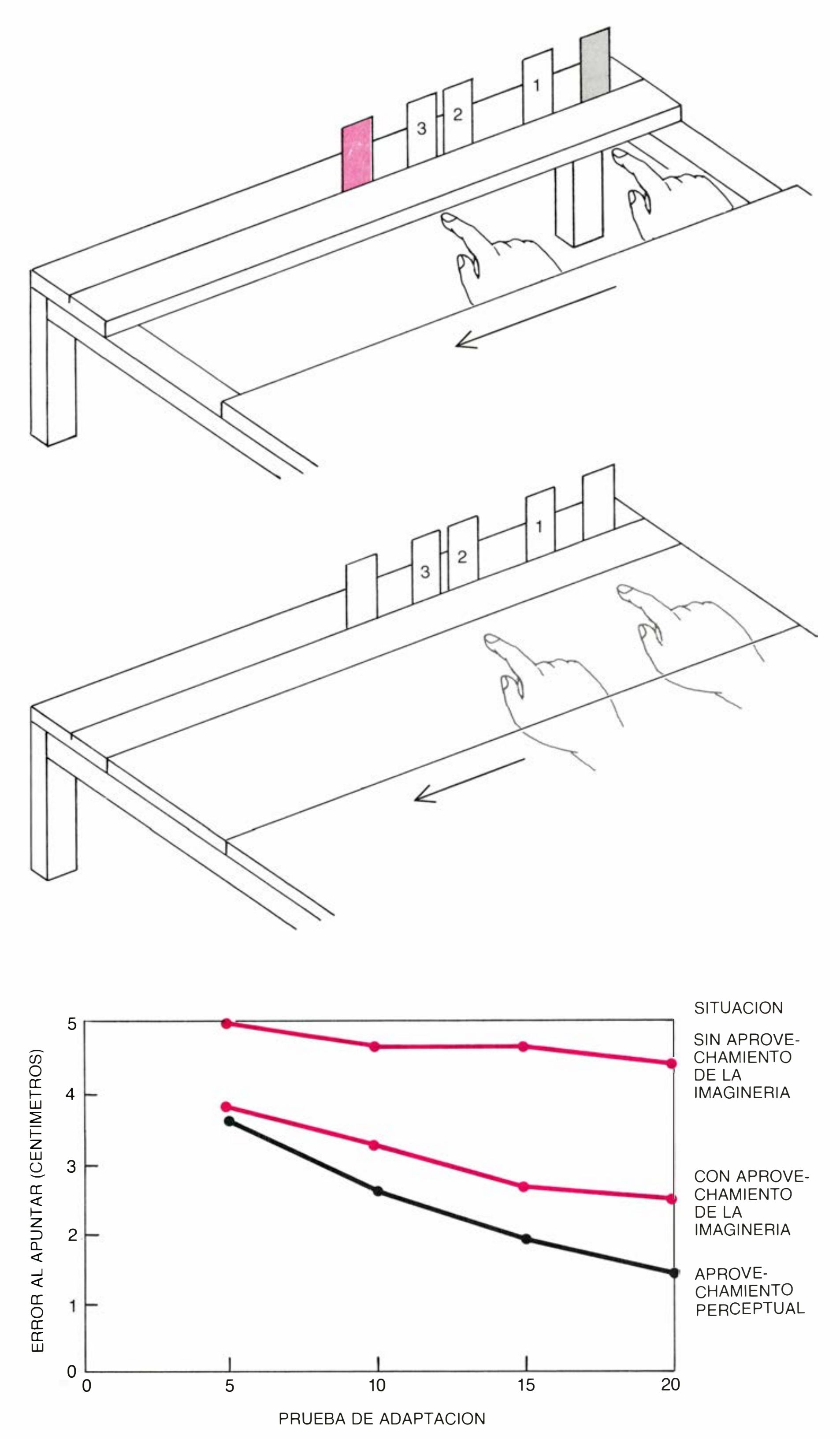
En mis experimentos los sujetos llevaban unas gafas especiales dotadas de prismas ópticos. A los de un grupo les pedía que extendieran el brazo derecho y apuntaran hacia un marcador rojo colocado frente a ellos sobre una mesa, a la altura de sus ojos. Por efecto de los prismas, al principio apuntaban unos cinco centímetros a la derecha del marcador. Puesto que los prismas desplazan el campo visual entero, una vez extendidos los brazos se percataban los sujetos de sus errores, que iban corrigiendo en los sucesivos intentos.

Medí los errores y calculé su promedio en grupos consecutivos de ensayos; expuse luego el promedio de localizaciones erróneas con tres marcadores. Los marcadores se utilizaron con un segundo grupo de sujetos a los que llamaré "sujetos imaginería". Llevaban también las gafas especiales, pero entre ellos y la mesa que sostenía los marcadores había un cartón que impedía pudieran verse los dedos al extender los brazos. Pedí a los sujetos imaginería que apuntaran hacia el marcador rojo mirando a través de los prismas y que imaginaran luego que tan pronto el brazo se extendía del todo, el dedo alcanzaba el correspondiente marcador de error.

En otras palabras, los marcadores de error aseguraban que los errores imaginados solían corresponder al promedio de errores cometidos por el primer grupo de sujetos. Añadí un tercer grupo a modo de control. Los individuos de éste apuntaban al marcador rojo sin beneficiarse de observar sus errores ni de que se les dijese que los imaginaran.

Sólo los sujetos puestos en condiciones de percepción y de imaginería mental mostraron una significativa reducción de errores al apuntar. Es más, sus tasas de adaptación fueron parecidas. Los resultados para los postefectos del apuntar, que se producen al retirarse las gafas, corroboran todavía más que existe una equivalencia funcional entre los errores observados y los imaginados. Aunque los postefectos en la situación de imaginería fueron menores que los que se dieron en la perceptual, los sujetos de ambos grupos apuntaron hacia la izquierda del marcador rojo cuando se restablecieron las condiciones de la visión normal. También obtuve evidencia de cambio intermanual: los sujetos no sólo apuntaban a la izquierda con su mano derecha (la mano "adaptada"), sino que apuntaban igualmente en ese sentido con la izquierda (la mano "inadaptada").

Son diversas las implicaciones de esos resultados. En primer lugar, resulta muy improbable que los sujetos pudieran predecir las características adaptativas y transferenciales de los cambios inducidos por los prismas en su coordinación visual-motora. Parece asimismo inverosímil que tuvieran cualquier tipo de experiencias visuales afines que les proporcionasen un conocimiento inconsciente de esos efectos. En segundo lugar, los resultados demuestran que la imaginería mental puede producir ciertos cambios de la coordinación visual-motora, que persisten después de que las imágenes no se formen ya. Sugieren también que la




**4. VALOR FUNCIONAL DE LA IMAGINERIA MENTAL. Se estimó considerando el papel que las imágenes pudieran desempeñar en la adaptación a un prisma. Los prismas ópticos desplazan la localización aparente de un objeto. A sujetos provistos de gafas portadoras de tales prismas se les mandó apuntar a un blanco (*marcador de color de la hilera superior*). En virtud del efecto de los prismas, los sujetos apuntaban al principio cinco centímetros a la derecha del marcador (*marcador gris*). Sin embargo, y puesto que los prismas desplazan el campo visual entero, al extender los brazos los sujetos experimentales advertían su error y lo corregían en los siguientes intentos (*marcadores 1-3*). A continuación se utilizaban los marcadores con un segundo grupo de sujetos, denominados "sujetos imaginería". Estos llevaban también las gafas especiales, pero el espacio que quedaba entre ellos y la mesa que sostenía los marcadores se cubría con un tablero, de modo que no pudieran verse los dedos (*segunda hilera*). Su tarea consistía en imaginar que su dedo apuntaba al correspondiente marcador del error en cuanto extendían el brazo. En un tercer grupo, el de control, los sujetos apuntaban hacia el marcador de color sin permitírseles observar sus errores y sin decirles que se los imaginaran. El gráfico resultante muestra que se seguía una significativa cantidad de reducción del error cuando los sujetos imaginaban sus errores, casi tanta como cuando de hecho los advertían.**

utilización de la imaginería mental para precipitar tales cambios pudiera tener importantes aplicaciones prácticas. Los atletas profesionales, por ejemplo, declaran a menudo que les resulta muy provechoso ensayar mentalmente su actuación; a la luz de esos experimentos es razonable pensar que el buen éxito de tales técnicas depende de la claridad y la exactitud con que se imagina la actuación.

Los frutos de las investigaciones que acabo de exponer ilustran de muchos modos que las imágenes mentales pueden corresponder funcionalmente a objetos y fenómenos físicamente percibidos. Cuestión de mayor importancia práctica es la de si la imaginería mental podrá facilitar directamente los procesos perceptuales en curso, suponiendo que se dé una equivalencia funcional. Puesto que hay impuestas similares constricciones en la imaginería y la percepción, ¿sería posible, por ejemplo, ver un objeto más rápidamente si se formase una imagen apropiada del mismo antes de su aparición real?

Según lo propusieron hace unos 10 años Ulric Neisser y Lynn A. Cooper, entonces en Cornell, y Roger N. Shepard, de la Universidad de Stanford, el proceso de formación de una imagen mental puede desempeñar una función anticipatoria de la percepción: puede preparar a la persona para recibir información sobre objetos imaginados. Tal vez, pues, la imaginería mental intensifique la percepción de un objeto aprestando selectivamente los apropiados mecanismos neurviosos del sistema visual. En otros términos, el que se forme la imagen mental de un objeto quizás inicie ciertos fenómenos nerviosos equivalentes a los que ocurren en el momento en que el objeto se ve, facilitando con ello el proceso perceptual.

Sin embargo, si aparece un objeto diferente del imaginado, puede que la formación de la imagen obstaculice la operación normal del sistema visual. Suponga usted, por ejemplo, que va volando en un avión a través de una nube. Podría ver antes la pista de aterrizaje si se la imaginara también antes de haberla localizado. Si, por otra parte, se la imaginara usted localizándola desacertadamente, le costaría más tiempo corregir el error que si no se la hubiese imaginado de ningún modo.

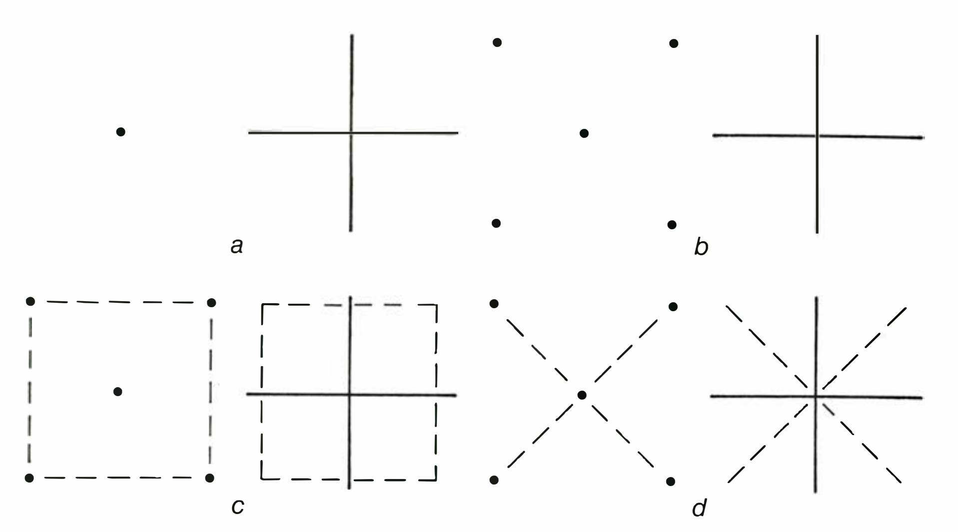
Algunos experimentos que realicé no hace mucho en Stanford ayudan a esclarecer la relación práctica que se establece entre formación de la imagen y percepción del objeto. En uno de los

ensayos los sujetos indicaban si sobre una pantalla circular veían una barra horizontal o una vertical. A esas dos barras alternativas se las llamaba “barras diana”. Les decía yo a los sujetos que formasen, de antemano, una imagen mental de una barra idénticamente configurada y orientada entre la vertical y la horizontal, o bien que no formasen ninguna imagen mental. En cada ensayo, pues, una de las dos barras diana alternativas se superponía a una barra imaginada (en los casos en que se imaginaba una barra). Registré el tiempo de reacción para la identificación de la barra diana en función del respectivo alineamiento de la misma con la barra imaginada. Los tiempos de reacción para los casos sin imagen servían de medida basilar.

Los sujetos identificaban antes las barras cuando las imaginadas se alineaban muy cerca de las diana, dentro de una escala de alrededor de 10 grados. Al incrementarse hasta 45 grados el ángulo entre las barras imaginadas y las diana, aumentaba también el tiempo necesario para la identificación. Para ángulos mayores de 45 grados disminuía de nuevo el tiempo. En otras palabras, la máxima interferencia en la identificación se producía cuando las barras se habían imaginado exactamente entre las dos posibles orientaciones de las barras diana.

¿Por qué no aumentaron los tiempos de reacción en proporción directa al grado de desalineamiento entre las barras imaginadas y las diana? Una razón de ello es que la selección de respuestas de los sujetos podría haberse basado en la comparación entre la imagen mental y la diana. Si la imagen se ajusta a la diana, la respuesta correspondiente a la orientación de la imagen se selecciona pronto. Si la imagen y la diana difieren en 90 grados, la comparación indica que la respuesta correcta es la contraria a la que corresponde a la orientación de la imagen. Si la barra imaginada está, empero, entre las orientaciones de las dos barras diana, la comparación provoca confusiones y la imagen estorba el proceso de decisión.

Un segundo experimento abona esa explicación. En él mandé a los sujetos que indicasen lo más rápidamente posible si aparecía alguna de las dos barras diana. En el ensayo previo tenían que distinguir una diana de la otra; ahora sólo debían detectar la presencia de alguna de ellas, sin necesidad de identificarla. Los resultados demostraron claramente que la imaginería mental, bajo esas condiciones, no afectaba en nada a los simples juicios de detección. Parece, por tanto, que si bien la formación de imagen pueda influir en la identificación de muestras visuales, no lo hace en los procesos, más elementales, de simple detección y cambio de estímulo.




**5. AYUDA LA IMAGINERIA MENTAL a las tareas de discriminación. Se pidió a los sujetos que determinaran si la longitud de una línea vertical era mayor que la de otra horizontal, o viceversa. En la situación *a*, los sujetos fijaban la mirada en un punto y se les mostraba entonces las dos líneas, centradas en ese punto. En *b* se les mostraba primero cuatro puntos alrededor del de fijación. En *c* se les presentaba los mismos puntos que en *b*, pero pidiéndoles que formaran la imagen mental de un cuadrado que los uniera. En *d* se les instaba a formar la imagen de un aspa, ×, que conectara los cuatro puntos. La máxima facilitación a la hora de emitir los juicios sobre la longitud de las líneas se obtuvo imaginándose de antemano el cuadrado.**

Experimentos ulteriores, realizados por Jennifer J. Freyd y el autor en Stanford, aportaron pruebas de un tipo de facilitación de la imagen que no se explica a partir de la selección de la respuesta. En esos experimentos estudiamos los efectos de formación de una imagen mental que pudieran servir de valioso o inútil contexto visual para efectuar discriminaciones difíciles entre longitudes. Presentábamos a nuestros sujetos figuras consistentes en dos líneas rectas y les pedíamos que indicaran cuál era más larga. Al comienzo de algunos ensayos decíamos a los sujetos que formasen una imagen de un cuadrado lineal, que si de hecho se superpusiese en el centro de la muestra de las dos líneas haría resaltar las pequeñas diferencias entre sus longitudes. Durante otras pruebas mandábamos a los sujetos formar una imagen de un aspa, ×, (cuyos puntos terminales correspondían a los cuatro ángulos del cuadrado imaginario), lo cual, era de presumir, no resultaría tan útil para establecer las discriminaciones.

Hallamos que la formación de una imagen mental de la figura contextual auxiliar (el cuadrado) reducía el tiempo necesario para discriminar las longitudes, en comparación con el tiempo requerido cuando la imaginada era la figura menos útil (la ×) o cuando no se mandaba a los sujetos formar imágenes mentales. Más aún, los efectos fueron similares a los obtenidos cuando se presentaban realmente las mismas figuras contextuales.

Comprobamos también que los sujetos decidían a menudo imaginarse la figura contextual auxiliar al presentárseles indicadores posicionales que les servían para imaginarse una figura u otra. Dado que las figuras contextuales mismas no podían haber influido en la selección de las dos respuestas alternativas, ese tipo de facilitación de la imagen no podía derivar de un proceso interno de igualación o de selección de la respuesta. Antes bien, quizá se debiera a una síntesis mental de rasgos reales y rasgos imaginarios efectuada en algún nivel superior del sistema visual, donde el añadido de información contextual puede que resalte las diferencias entre los objetos sometidos a comparación.

En los procedimientos hasta aquí descritos se les pedía explícitamente a los sujetos experimentales que formaran algún tipo de imagen mental. Comporta ese proceder una posible falacia: quizás anime a los sujetos a tratar de actuar como lo harían en una tarea perceptual correspondiente, pensando que eso es lo que se supone que deben hacer. Aunque el problema puede evitarse en gran parte procurando medir efectos perceptuales sutiles o inesperados, un recurso todavía mejor es el de evidenciar que pueden formarse imágenes mentales para algún propósito específico, aun cuando no se haya

dado ningún tipo de instrucciones sobre la imaginería.

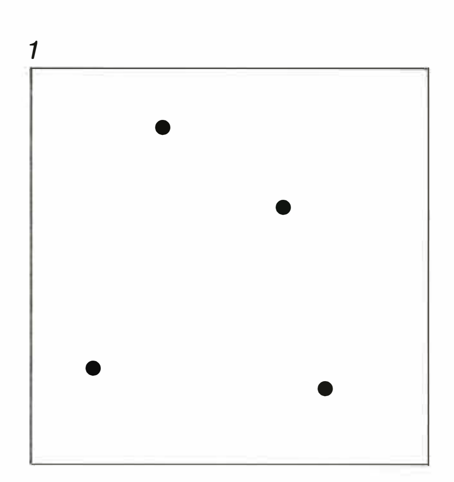
Estas importantes consideraciones son consecuencia de los primeros estudios efectuados por Kosslyn y Pinker sobre el escrutamiento de la imagen mental. Pedían éstos a sus sujetos que examinasen configuraciones de objetos (tales como las señales con que se trazan en un mapa las fronteras), que formasen una imagen mental de la figura y "fijasen la mirada" sobre uno de los objetos. Los investigadores nombraban entonces un segundo objeto e invitaban a los sujetos a ir escrutando mentalmente una senda que fuese directa, desde el primer objeto hasta el segundo. Kosslyn y Pinker hallaron en todos los casos que el tiempo que se requería para completar el escrutamiento de la imagen mental era directamente proporcional a la distancia física original entre los objetos, y concluyeron por tanto que las imágenes mentales conservan las características espaciales de una representación física.

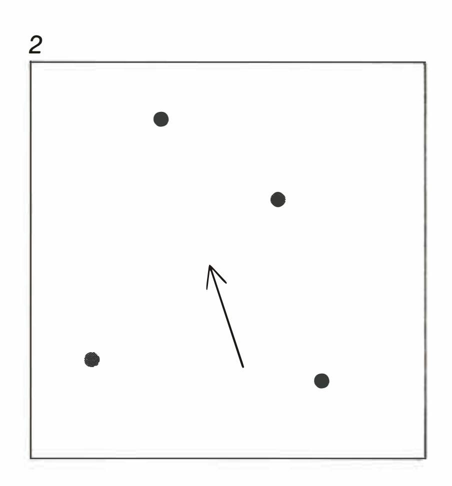
Se han criticado sus hallazgos porque a los sujetos experimentales no les sería difícil figurarse que las distancias mayores habrían de requerir tiempos de escrutamiento más largos. Pinker y yo desarrollamos luego un ejercicio que parece evitar este problema, pidiendo a un sujeto que forme imágenes mentales y las escrute sin explicitarle las direcciones en que ha de hacerlo. Tras haber examinado la figura de un punto, se les mostraba a nuestros sujetos una flecha y se les pedía que indicasen si ésta apuntaba a alguno de los puntos con anterioridad. Habíamos predicho que, para comprobar si se interceptaba alguno de los puntos, los sujetos tendrían que escrutar una imagen mental de la muestra siguiendo la dirección señalada por la flecha.

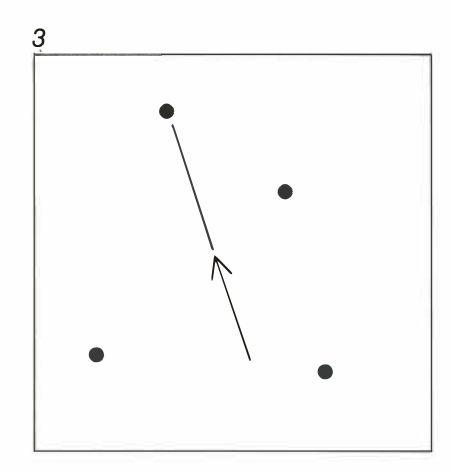
El experimento tuvo pleno éxito. Los tiempos de decisión aumentaron linealmente a medida que crecía la senda por examinar entre las flechas y los puntos. Es más, casi todos los sujetos declararon que, para cumplir la tarea encomendada, tuvieron que formar y escrutar una imagen mental del punto patrón. Demostramos de esa manera que el escrutamiento de la imagen mental puede resultar útil siempre que haga falta anticipar las consecuencias del irse moviendo a lo largo de una determinada senda desde un punto de partida dado.

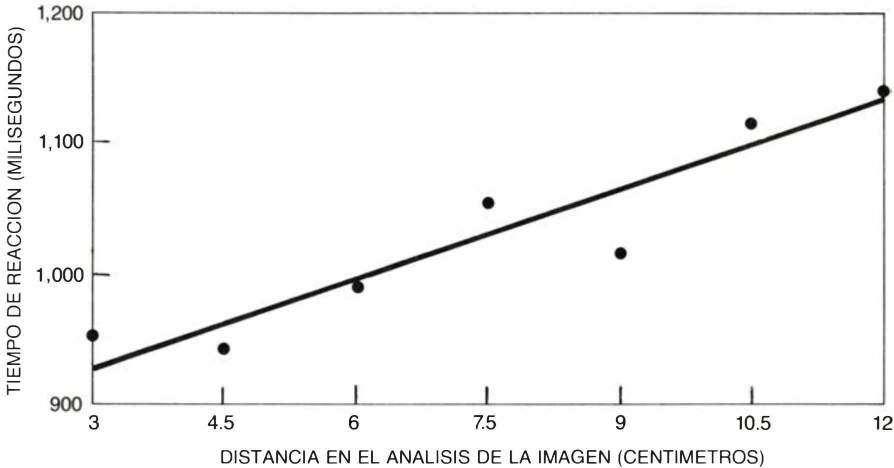
Supóngase que tratamos de imaginar dónde se detendría una bola de billar si la lanzáramos en determinada dirección. Aunque no se pudiera hacer rodar realmente la bola por la mesa, ni determinar matemáticamente la respuesta, sí que cabría figurarse lo que ocurriría resiguiendo mentalmente su movimiento y sus rebotes en las bandas. Cooper y Shepard han dado cuenta de resultados afines en las consecuencias imaginadas de la rotación de objetos [véase "Rotación mental de los objetos", por Lynn A. Cooper y Roger N. Shepard; INVESTIGACIÓN Y CIENCIA, febrero de 1985].

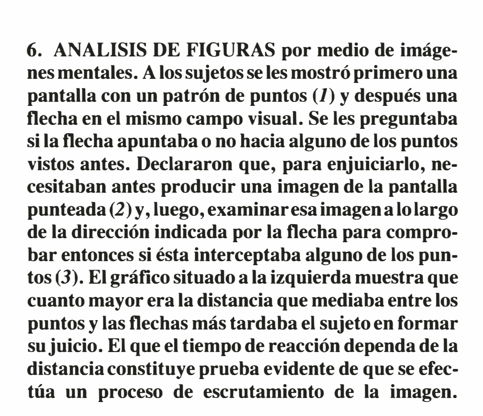
A la luz de estos estudios parece razonable proponer que siempre que la imaginería y la percepción comparten mecanismos nerviosos del sistema visual, podrían las imágenes mentales facilitar los procesos perceptivos que aquellos mecanismos sustentan. Debería intentarse, pues, determinar los más bajos niveles visuales en los que se comparten tales mecanismos. Si se concibe, por ejemplo, que la percepción de una muestra visual implica una ordenada secuencia de estadios de procesamiento de información, que abarca desde los más bajos hasta los más altos niveles del sistema visual, podría empezarse por tratar de descubrir cuán abajo de esa secuencia puede influir la



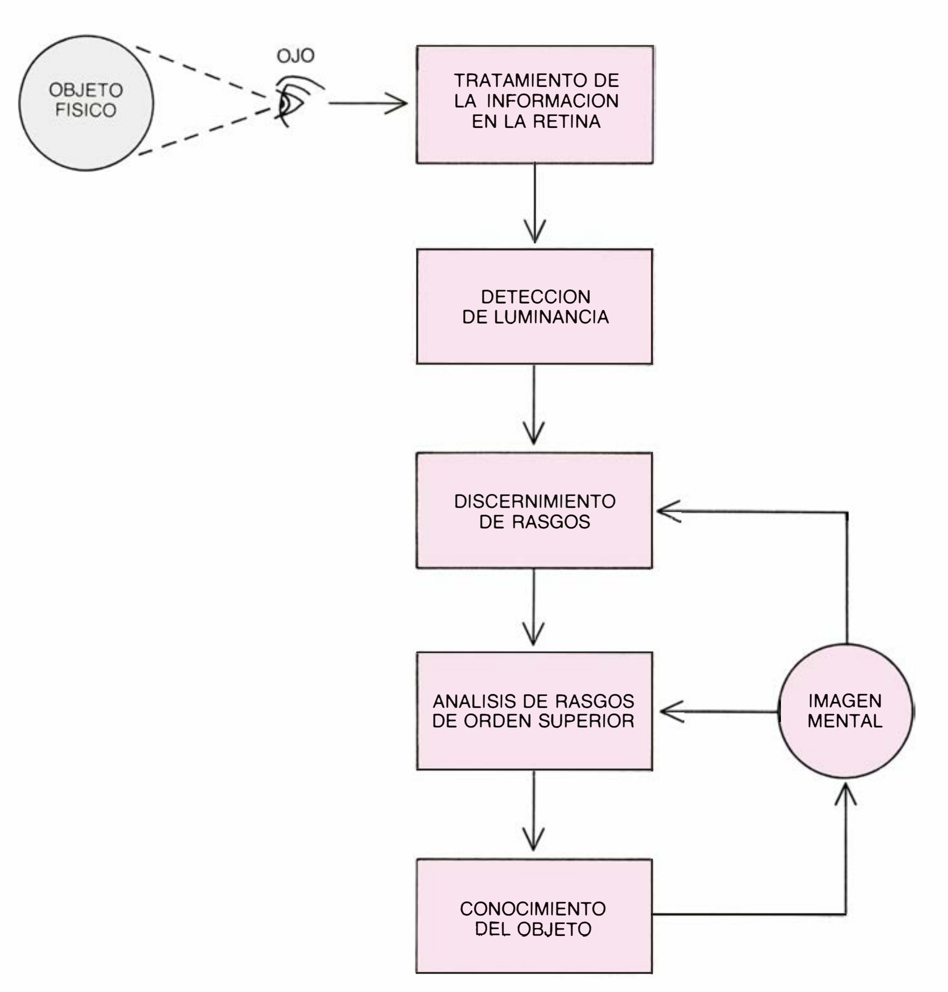



**6. ANALISIS DE FIGURAS por medio de imágenes mentales. A los sujetos se les mostró primero una pantalla con un patrón de puntos (*1*) y después una flecha en el mismo campo visual. Se les preguntaba si la flecha apuntaba o no hacia alguno de los puntos vistos antes. Declararon que, para enjuiciarlo, necesitaban antes producir una imagen de la pantalla punteada (*2*) y, luego, examinar esa imagen a lo largo de la dirección indicada por la flecha para comprobar entonces si ésta interceptaba alguno de los puntos (*3*). El gráfico situado a la izquierda muestra que cuanto mayor era la distancia que mediaba entre los puntos y las flechas más tardaba el sujeto en formar su juicio. El que el tiempo de reacción dependa de la distancia constituye prueba evidente de que se efectúa un proceso de escrutamiento de la imagen.**




**7. MODELO de cómo la imaginería mental podría influir en la percepción visual; se basa en los trabajos del autor y de otros investigadores. La percepción de un objeto es consecuencia de la activación nerviosa a lo largo de una serie de estadios de tratamiento de la información en el sistema visual, comenzándose en el nivel retiniano. La formación de una imagen del objeto está determinada por el conocimiento que la persona tiene de los rasgos característicos de ese objeto, y se produce presumiblemente en los niveles más altos. Una vez formada, una imagen puede afectar a mecanismos nerviosos de los niveles visuales intermedios, que son los responsables del discernimiento de rasgos y de otros tipos de análisis más complicados, modificando así, quizás, la percepción del objeto. La imaginería mental probablemente no influya en los niveles visuales inferiores a los que se ocupan del discernimiento de rasgos. (Los dibujos son obra de Jerome Kuhl.)**

formación de la imagen en los mecanismos subyacentes.

No es de creer que la imaginería mental ejerza un efecto excesivo en el nivel ínfimo, el retiniano, donde se hallan los tipos más primitivos de mecanismos de tratamiento de la información. Como tampoco que la imaginería mental altere el procesamiento de información en los niveles precorticales, donde a los mecanismos les corresponde la detección de los cambios de intensidad o los contrastes luminosos. Sólo en niveles algo más altos, responsables del discernimiento de formas (como ocurre en el córtex visual), empieza a advertirse que las imágenes mentales influyen sobre la percepción. En niveles superiores, las pruebas de que la imaginería influye sobre la percepción son contundentes.

Finalmente, cabe suponer que en los niveles supremos los procesos perceptuales cooperan con procesos más abstractos e intervienen en el conocimiento y la comprensión de los objetos físicos. Conviene aquí distinguir entre la forma y la función de una imagen mental. Cuando alguien decide crear una imagen mental de un objeto particular, el tipo de imagen que puede formarse depende del conocimiento que esa persona tenga del objeto, así su tamaño, color y figura. Después, una vez formada ya la imagen, puede ésta llegar a actuar en algunos aspectos como el objeto mismo, suscitando la activación de ciertos tipos de mecanismos nerviosos en los niveles inferiores del sistema visual. Por consiguiente, cualesquiera sean las restricciones que tales mecanismos impongan a la calidad de la percepción del objeto pesarán también sobre la calidad de la imaginería mental elaborada. Así, las imágenes mentales pueden adquirir características visuales y pueden servir para modificar la percepción misma.

# Almacenaje de grano en Assiros

*En este yacimiento del norte de Grecia, los restos de grano carbonizado, procedentes de almacenes sepultados hace 3000 años, arrojan nueva luz sobre la construcción de los majestuosos palacios micénicos de la Edad del Bronce*

Glynis Jones, Kenneth Wardle, Paul Halstead y Diana Wardle

Las sociedades de la Edad del Bronce del Egeo, incluidas la cultura minoica de Creta y la micénica de Grecia continental, fueron ricas y complejas. Ocupaba su centro el palacio, residencia del gobierno y foco de la acción religiosa. El palacio desempeñaba también otra función crucial: almacenaba el excedente agrícola. La autoridad de los soberanos de la Edad del Bronce descansaba, en buena medida, en su disposición de provisiones de vino, grano, aceite y otros productos de la región circundante. Los convenios económicos y sociales que subyacen tras un sistema tan elaborado de almacenaje atravesaron, sin duda, un largo proceso de desarrollo, historia que no podemos tejer todavía con seguridad. Las pruebas directas –los granos recogidos– son perecederas. Añádase que los objetos de arcilla, metal o piedra atrajeron tradicionalmente mayor interés que los restos vegetales en los estudiosos de la Edad del Bronce. De ahí que las formas de almacenaje de alimentos que condujeron a la economía palacial disten de conocerse bien.

Pero la situación empieza a cambiar. Los arqueólogos de la Edad del Bronce del Egeo han adquirido mayor perspicacia en lo concerniente a las lecciones a sacar del registro botánico. En esa línea, cuando excavamos un poblado de la Edad del Bronce en Assiros (Grecia septentrional), a comienzos de 1975, nuestro equipo incluyó un arqueobotánico (Jones) durante la campaña. Los resultados justificaron con creces su incorporación. Cuidadosos muestreos de semillas, tomadas de los suelos de los edificios que habían sido destruidos por el fuego en la Edad del Bronce, revelaban que algunas estancias habían estado destinadas al almacenaje de granos. Por lo que se vio, dicho almacenaje debió ser, en Assiros, de tipo comunal, adoptando una forma intermedia entre la despensa familiar y la recaudación organizada de excedente, propia de la economía palacial plena. Assiros pudo constituir una suerte de comunidad que desembocaría, bajo condiciones apropiadas, en un centro palacial. Lo que explica la importancia que dicha localidad adquiere para desentrañar los orígenes de los grandes palacios de la Edad del Bronce.

Assiros se encuentra en Macedonia. Aunque hoy día Macedonia forma parte de Grecia, durante la Edad del Bronce se situaba en la periferia del mundo micénico. La sociedad micénica estaba centrada en la península del Egeo: su límite septentrional corría al sur de Macedonia, a través de Tesalia. Macedonia era a la sazón una encrucijada cultural, abierta a influencias de la región danubiana por el norte, de Tracia por el este y de Albania por el oeste, así como del mundo micénico. De estas múltiples influencias, la del mundo micénico era una de las más fuertes. Está claro, sin embargo, que el desarrollo social de los grupos norteños era menos avanzado que el de sus vecinos meridionales de Micenas y Pilos en el continente y Cnossos en Creta. Nada de lo encontrado en Macedonia puede compararse con los grandes palacios desenterrados en aquellos y otros yacimientos [véase "Palacios minoicos", por Peter M. Warren; INVESTIGACIÓN Y CIENCIA, septiembre de 1985]. Este atraso relativo puede resultar de bastante ayuda para los arqueólogos, porque implica que los hallazgos de Assiros, contemporáneos de los palacios, pueden proporcionar información acerca de los requisitos previos para la constitución de la economía palacial.

Una razón por la que Assiros se nos ofrece como fuente potencial de información sobre los antecedentes de la economía palacial se apoya en su período de ocupación: coincide con la edad de los palacios micénicos. Estos últimos conocieron su apogeo entre 1400 y 1200 a.C. De acuerdo con nuestra investigación, la Assiros desenterrada estuvo permanentemente habitada durante un milenio que abarcaba esas fechas.

El yacimiento de la Edad del Bronce que excavamos es una elevación oval llamada Assiros Toumba, no lejos de la actual Assiros. El montículo alcanza unos 14 metros de altura. Se trata de un accidente completamente artificial. Los primeros habitantes llegaron alrededor del 1800 a.C. y construyeron un muro defensivo que ceñía las casas sobre lo que entonces era más o menos un terreno llano. Durante los siguientes 1000 años el muro y las estructuras interiores se reconstruyeron muchas veces. Con cada reconstrucción el montículo crecía y, como sus bordes se inclinaban hacia adentro, el área habitable intramuros decreció. En torno al 800 a.C., a comienzos de la Edad del Hierro, se abandonó el lugar, quizá porque el espacio encerrado, empequeñecido, no servía ya para acomodar a la población del asentamiento.

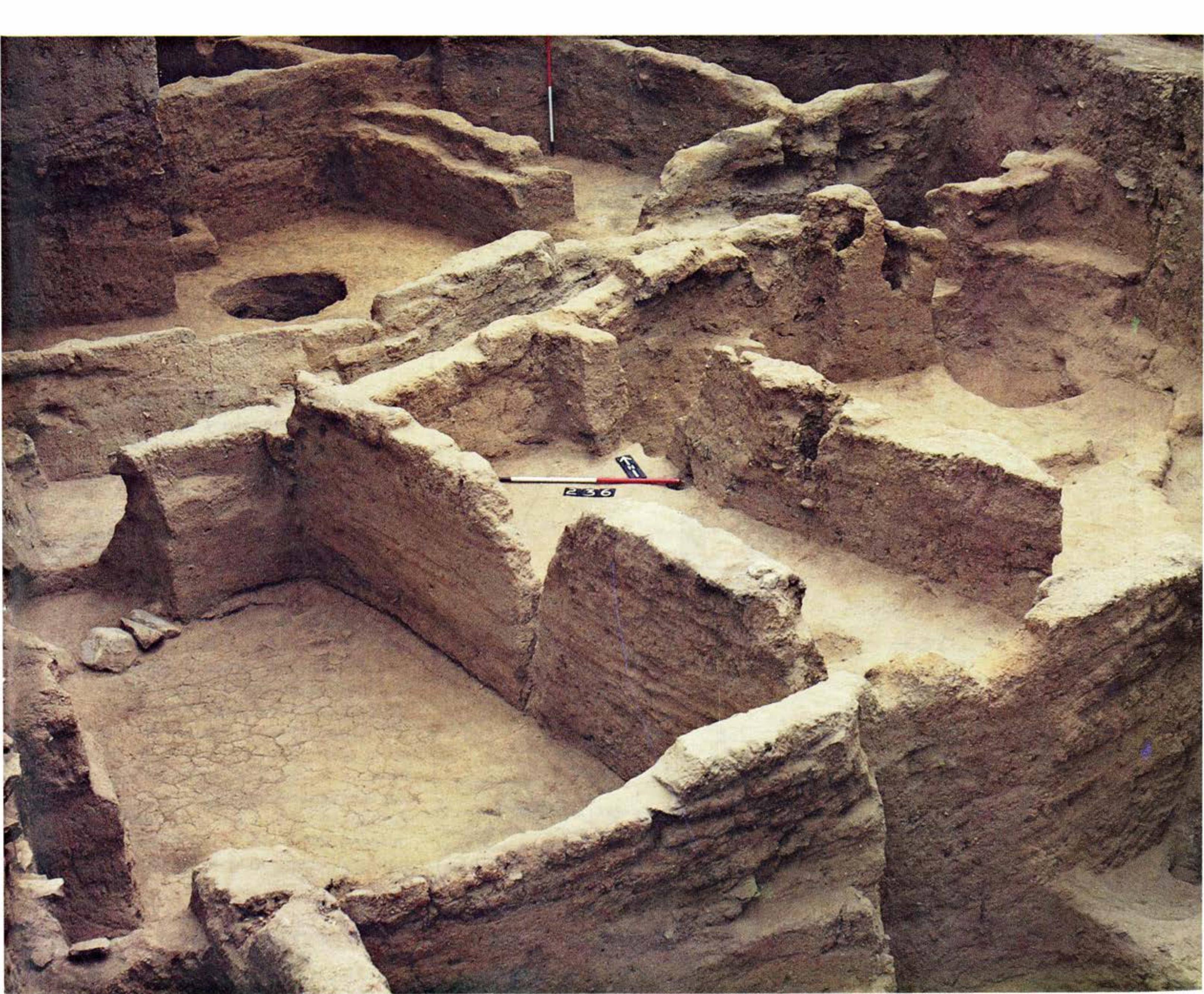
A pesar de sus reconstrucciones periódicas, los métodos de edificación parecen haber permanecido más o menos invariables a lo largo de casi toda la Edad del Bronce. Y así, el poblado levantado sobre el montículo debió cambiar poco en cada reconstrucción sucesiva. El muro circundante era de ar-

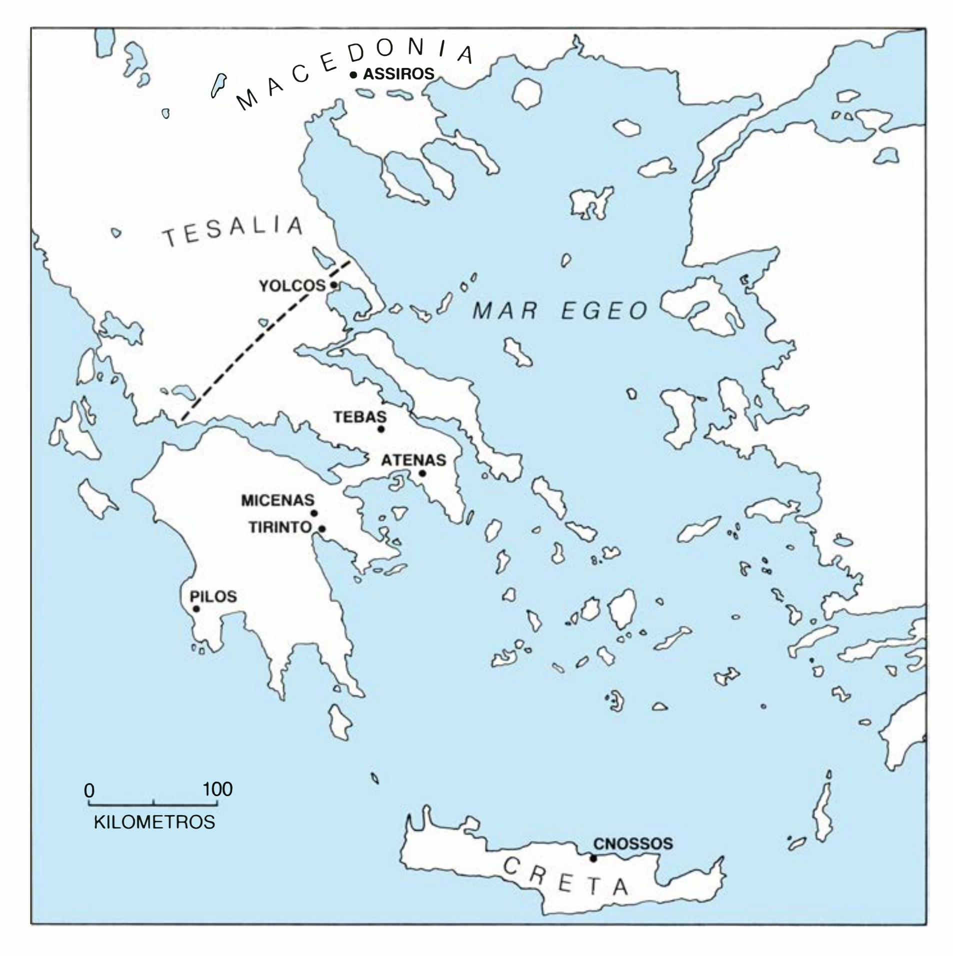
**1. ESTRUCTURAS DE LA EDAD DEL BRONCE, descubiertas por los autores en un gran montículo cerca del actual pueblo griego de Assiros. Los muros tienen armazones de madera que están rellenos con ladrillos de adobe secado al sol. Los muros formaron un panal de habitaciones y almacenes separados por patios y callejones. Un muro de ladrillos de adobe ceñía el asentamiento. Las habitaciones de la fotografía se ocuparon alrededor del 1200 a.C., al final de la Edad del Bronce. Bajo ellas están los suelos de los almacenes de grano que fueron destruidos por un incendio alrededor del 1350 a.C. Una gran cantidad de restos botánicos se ha encontrado entre los derrumbes carbonizados sobre los suelos. El pozo circular del fondo era para ensilaje; las depresiones de la derecha de la fotografía son resultado de remociones tardías de la Edad del Hierro.**

cilla y ladrillos de barro secado al sol. En su interior se erigían las habitaciones, rectangulares, dotadas de armazones de madera rellenos con ladrillos de adobe. (Los armazones de madera pretendían, tal vez, reforzar las viviendas contra los terremotos que asolaron Macedonia durante la Edad del Bronce y continúan haciéndolo todavía.) Entre las habitaciones, que sirvieron de viviendas y para otras funciones, había patios. Grupos de habitaciones y patios estaban divididos por callejones, confiriendo al asentamiento un esquema rectilíneo. Los callejones, cuyo piso era de grava, resultaban demasiado estrechos para permitir el paso de un animal de carga y menos aún el de un carro; hay que pensar, pues, que estaban hechos fundamentalmente para caminar.

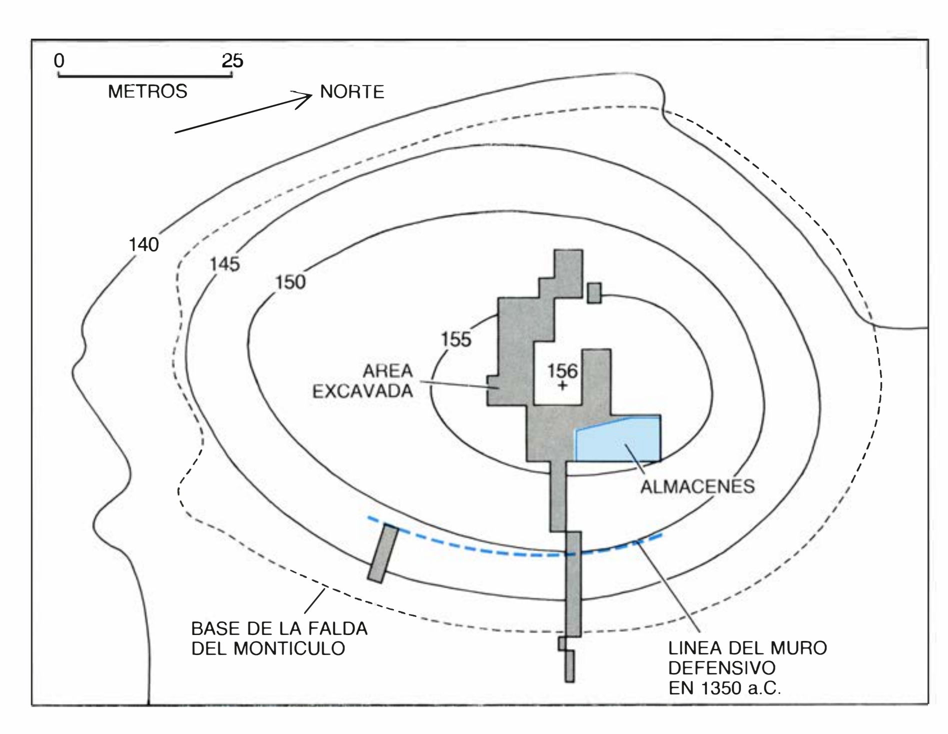
En la parte septentrional de este panal de habitaciones, patios y callejones descubrimos los almacenes. El hallazgo, el más estimulante de cuantos logramos, nos llegó de todo punto inesperado. Al iniciar la excavación, suponíase que Assiros Toumba ocultaba un yacimiento de la Edad del Bronce, pero se desconocía la naturaleza del asentamiento y la duración de su ocupación. El plan original consistía en excavar un 10 por ciento, aproximadamente, del montículo hasta una profundidad de tres metros con el fin de documentar la secuencia de los niveles de ocupación. En el curso de esa tarea, tomamos muestras de restos de plantas de forma sistemática en cada nivel. En 1977, acabada la tercera campaña de excavación, el estudio de los restos botánicos en el laboratorio mostró que una zona del extremo norte aportaba una cantidad abundante de restos vegetales. En la siguiente campaña concentramos nuestro trabajo allí; para sorpresa nuestra, una afortunada coincidencia había preservado una colección intrigante de restos relacionados con la economía agrícola del asentamiento de la Edad del Bronce.

Sucedió que, alrededor del 1350 a.C., un incendio había destruido varias habitaciones del poblado. Dispersos por el suelo de las habitaciones quemadas había ladrillos de adobe, quemados también, y restos carbonizados de la madera de la techumbre. Las gentes de Assiros no se habían preocupado de retirar los escombros y reconstruir los suelos de los edificios incendiados. Limitáronse a nivelar las ruinas y convertir la nueva superficie plana en basamento de un nuevo piso. Los armazones de madera de los edificios quemados fueron cortados casi por la base y los tocones se convirtieron en fundamentos de las nuevas estancias. Como consecuencia de este método de construcción "nuevo sobre viejo", una cantidad importante de grano quedó entre los escombros sobre el antiguo suelo. Restos de grano carbonizado restaron dispersos entre los ladrillos de adobe caídos. En particular, quedó grano concentrado en un nivel sobre el piso antiguo; en algunos puntos, alcanzó incluso varios centímetros de po-






**2. ASSIROS Y EL MUNDO MICENICO mantuvieron una estrecha relación durante el período de los palacios de la Edad del Bronce. Assiros se encuentra en Macedonia, que hoy forma parte de la Grecia septentrional; en el apogeo de los palacios (del 1400 al 1200 a.C.) constituía un área cultural autónoma; la frontera norteña del mundo micénico (*línea discontinua*) quedaba al sur de Macedonia, en Tesalia. Había palacios micénicos en Micenas, Tirinto, Pilos y Tebas y quizás en Yolkos y Atenas, aunque los datos son menos seguros.**

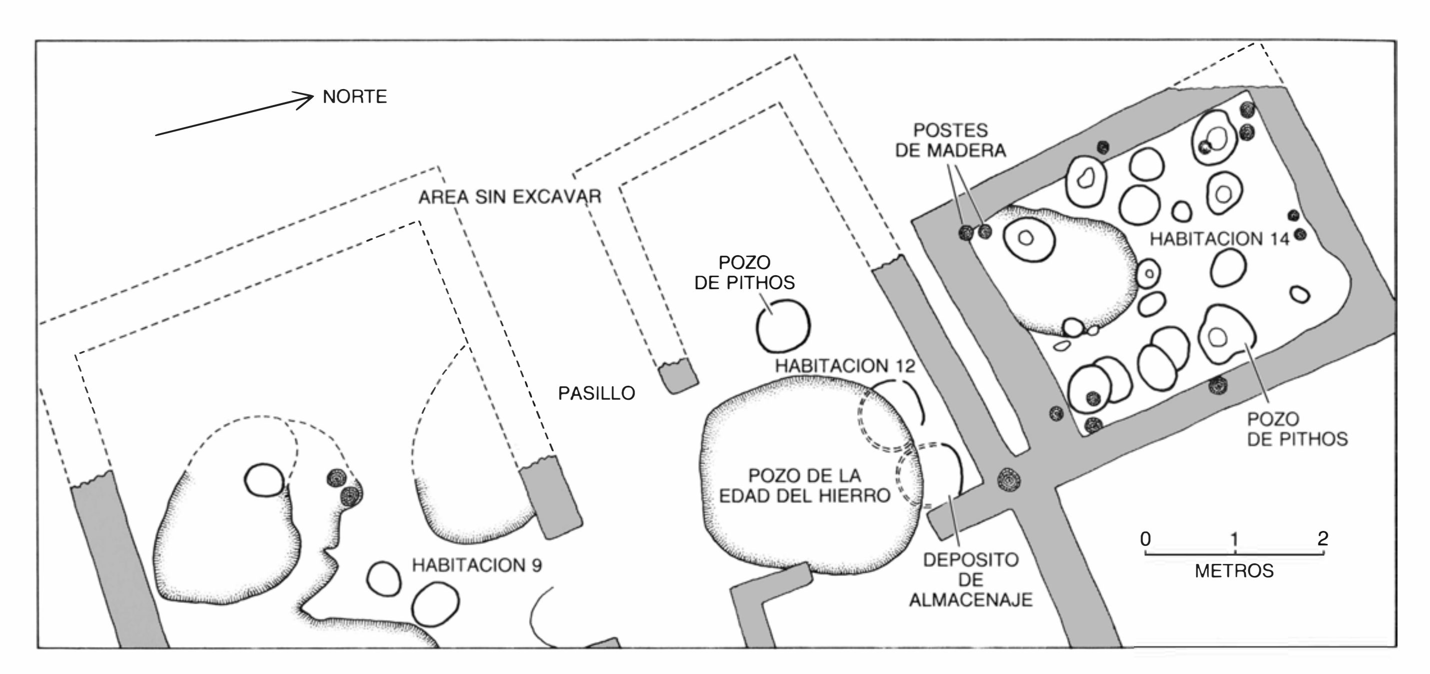



**3. ASSIROS TOUMBA es el montículo donde se hallaron los restos del poblado de la Edad del Bronce. El montículo, con unos 14 metros de altura, se formó en virtud de la construcción y reconstrucción del asentamiento, 20 veces o más, entre 1800 a.C. (a mediados de la Edad del Bronce) y 800 a.C. (en la Primera Edad del Hierro). El área excavada esta señalada en gris y los almacenes en color. (Esquema de Tom Prentiss.)**

tencia. Tres habitaciones (denominadas 9, 12 y 14) contenían gruesos niveles de restos carbonizados, así como también un pasadizo cubierto adosado. Una cuarta habitación y un patio adyacente parecen haber formado parte de la misma unidad arquitectónica, pero el incendio no los destruyó y, por tanto, no proporcionaron la misma riqueza de restos botánicos.

El material vegetal y de otra índole hallado en el complejo dejó pocas dudas: tratábase de un conjunto de almacenes de grano. Entre los restos no botánicos había indicios de objetos asociados con el ensilaje. En el suelo de la habitación 9, seis hoyos, y quizá más, habían sido excavados. Los hoyos, de unos 30 centímetros de profundidad, encajan exactamente con las bases apuntaladas de recipientes cerámicos encontrados en otros puntos del montículo. Tales recipientes, que eran las vasijas comunes de almacenaje de la Edad del Bronce, son conocidos en Grecia como *pithoi* (*pithos* en singular). El suelo de la habitación 14 estaba mejor conservado y poseía una decena larga de *pithoi.* Aunque en otra encontramos una pieza de un disco de arcilla, que podría haber sido la tapadera de un *pithos,* se hallaron pocos fragmentos de las vasijas. Parece que los recipientes de almacenaje no fueron dañados por el fuego y se salvaron. Gran parte del suelo de la habitación 12 había sufrido remociones a lo largo de la Edad del Hierro y no ofrecía señales de depresiones para las vasijas de ensilaje. Esa habitación, sin embargo, contenía dos depósitos de almacenaje cuyo armazón de mimbre se había recubierto con arcilla.

Tras ponderar los datos disponibles sobre el número y tipo de contenedores, calculamos la capacidad de almacenaje del complejo, o al menos de las tres habitaciones donde se encontraron los restos quemados. Parece que las habitaciones 9, 12 y 14 se dedicaron, de forma exclusiva, para almacenaje y que el grano se introdujo en *pithoi,* depósitos de arcilla y quizá, también, en sacos u otro tipo de contenedores perecederos que no se han conservado. Al hacer nuestro cómputo, partimos de la hipótesis de que el emplazamiento de los *pithoi* en el suelo de cada almacén siguió el patrón espacial observado en la parte intacta del suelo de la habitación 9, donde se hallaron seis pozos. De acuerdo con ese supuesto, y teniendo en cuenta que cada *pithos* albergaría entre 100 y 150 litros, calculamos una capacidad de almacenaje de




**4. EL COMPLEJO DE ALMACENAJE incluía tres habitaciones que fueron destruidas por el fuego. En los suelos había hoyuelos, cada uno de los cuales habría afianzado un *pithos*. La habitación 12 contenía restos de dos depósitos de almacenaje, que tenían un armazón de mimbre recubierto por arcilla. Después del incendio, los habitantes salvaron los recipientes, dejando el grano. Los restos quemados se nivelaron entonces para formar la base de un nuevo basamento. En algunos lugares, el grano alcanzó varios centímetros de espesor sobre el antiguo suelo, proporcionando unas muestras excelentes de las cosechas.**

unos 1000 litros para las habitaciones 12 y 14, cada una, y de 2000 a 3000 litros para la habitación 9. El volumen total, entre 4000 y 5000 litros, es más que suficiente para alimentar a 20 personas durante un año. La concentración de ese volumen de ensilaje en un área del montículo resulta notable para la Grecia prehistórica, donde la norma era la dispersión de los depósitos en cada casa. El único paralelo relativo a la disposición concentrada vista en Assiros ha de buscarse en los palacios de la Edad del Bronce del sur de Grecia.

En el pasado, el arqueólogo que hubiera encontrado almacenes como éstos probablemente hubiera creído que el grano así disperso habría sufrido tan profunda alteración durante y después del incendio, que difícilmente aportaría información segura sobre la forma original de almacenaje. En Assiros, sin embargo, un muestreo detallado nos permitió deducir muchas cosas sobre el número de especies de grano y su disposición espacial. La mayor parte de la información se obtuvo en el suelo antiguo que queda por debajo del nivel de las ruinas carbonizadas. Aunque el grano mezclado con los escombros padeció una intensa alteración, el que quedó en el suelo parecía bastante entero. En muchas zonas el grano disperso incluía una proporción muy alta de semillas de una sola especie vegetal. Dedujimos que aquellas concentraciones de una especie debieron permanecer inalteradas después del incendio. Cualquier transformación importante, sin duda, habría mezclado los contenidos de varios recipientes, resultando así combinaciones aleatorias de distintas especies y no concentraciones de un mismo tipo. Inferimos que esas concentraciones encerraban una información significativa sobre el sistema de ensilaje. Se analizó una por una. Se tomaron muestras del centro y la periferia de la concentración, así como de las zonas entre dos concentraciones. Si una concentración era grande se tomaban muchas muestras.

El muestreo botánico reveló, en primer lugar, que los habitantes de Assiros habían cultivado, al menos, siete especies de plantas. Tres especies de trigo fueron perfectamente identificadas: esprilla (*Triticum monococcum*), escanda (*T. dicoccum*) y espelta (*T. spelta*). Las semillas de otras dos especies –trigo candeal (*T. aestivum*) y trigo duro (*T. durum*)– son indiferenciables. La presencia de ese tipo de granos nos faculta para pensar en el cultivo de ambas especies; pudiera ocurrir, no obstante, que sólo se sembrara una. Además del trigo cosecharon, por lo menos, otros dos cereales: la cebada de seis carreras (*Hordeum vulgare*) y el mijo (*Panicum miliaceum*). Había también una legumbre llamada arveja (*Vicia ervidia*). Las siete especies constituyeron, probablemente, el núcleo de la dieta de los habitantes de la Edad del Bronce. Los cereales pudieron haberse consumido panificados o como gachas.

El potencial de un muestreo botánico espacialmente intensivo del estilo del acometido en Assiros no se limita a confeccionar una lista de las especies halladas. Entre otras cosas, revelará si los granos se cultivaron y almacenaron por separado. De acuerdo con las muestras, casi puras, de esprilla, mijo y arveja, dedúcese que estas especies fueron almacenadas –y por tanto presumiblemente sembradas– como cultivos separados. El aplastante predominio de trigo candeal, trigo duro o cebada en algunas muestras abona la hipótesis de un cultivo separado para cada una de estas especies. El caso de la esprilla y la espelta es más complejo. Aunque la proporción de las dos especies varía de una muestra a otra, siempre, sin embargo, aparecen juntas: lo lógico es aceptar que esprilla y espelta se cosecharon y ensilaron conjuntamente. En este sentido recuerda el "maslin", o cosechas mixtas de trigo y cebada que siembran hoy algunos campesinos griegos. Puesto que las dos plantas crecen en condiciones bastante diferentes, combinándolas se reduce la posibilidad de una pérdida total de la cosecha.

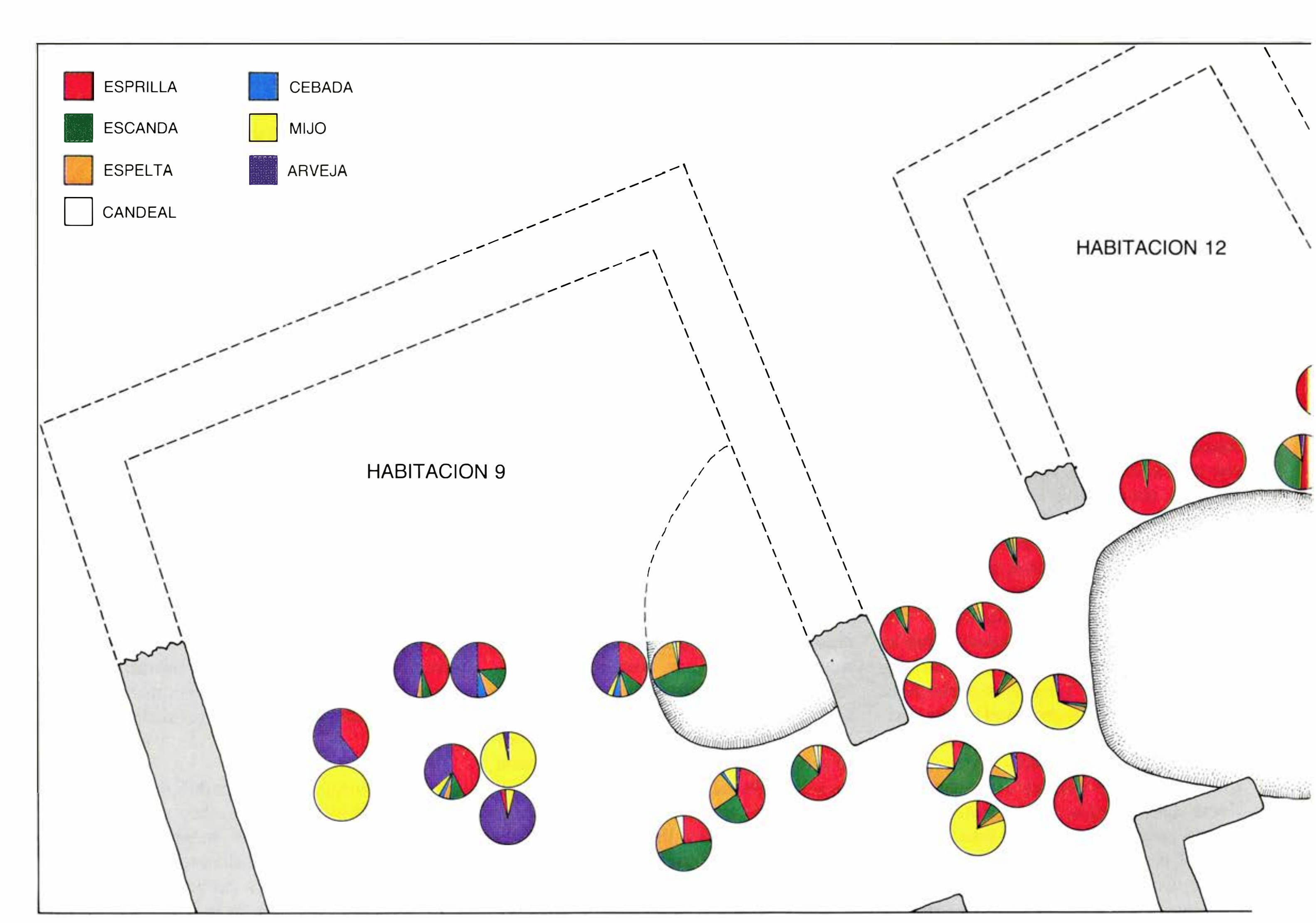
En su conjunto, la diversidad de granos cultivados en Assiros supondría también un factor beneficioso. Cada especie de las seis documentadas (esprilla, mijo, arveja, candeal, trigo duro y cebada, además de la mezcla de esprilla y espelta) se siembra y cosecha de forma diferente. Por tanto, cultivando esa gama de granos se extienden los períodos de labor agrícola a lo largo

de todo el año. Además, cada especie presenta exigencias particulares en lo relativo al clima y precipitaciones; el cultivo de varias puede servir de pantalla cautelar ante la imprevisible evolución de la temperatura y la pluviosidad. Las lluvias, sobre todo, constituirían la principal incógnita para las gentes de Assiros, ya que la parte oriental de Grecia tiene un clima árido en el que las precipitaciones no siguen ninguna pauta regular.

Las seis especies se sembrarían en ciclos diferentes y tendrían su forma propia de ensilaje. Gordon C. Hillman, de la Universidad de Londres, ha estudiado el procesado de los granos en los actuales asentamientos agrícolas primarios (entre éstos, la aldea turca). Trátase de economías similares a las vigentes en la Grecia de la Edad del Bronce. El trabajo de Hillman nos ayudará, pues, a interpretar ese procesado en tiempos prehistóricos. Sostiene que algunas plantas se sometían a un procesamiento completo antes de almacenarse la parte comestible. En la antigua Assiros, ese grupo incluiría: arveja, cebada, mijo, trigo candeal y trigo duro. Aquí sólo se requiere una trilla simple para separar las semillas del resto de la planta, seguido de un aventado y un tamizado con que recuperar el grano. A los depósitos llega el grano ya preparado para cocinar.

Otro grupo de plantas necesita dos fases en su procesado. La cosecha puede almacenarse al final de cualquiera de las dos. Entran aquí la escanda, la esprilla y la espelta, cuyas semillas están recubiertas herméticamente por gruesas vainas llamadas glumas. (De ahí su nombre de trigos de gluma.) El grano y su cascarilla constituyen la espiguilla; la espiga o mazorca del grano consta de muchas espiguillas adheridas al tallo central. Para que la cosecha se transforme en alimento, hay que trillar primero la siega para romper la mazorca en espiguillas individuales. La mezcla es entonces aventada para recuperar las espiguillas, dejando aparte el resto de la planta no comestible. Las espiguillas sufren un segundo trillado para separar el grano de la paja, formada principalmente por las vainas. El grano se recupera bien por un segundo aventado o por un tamizado.

El proceso en dos fases da lugar a dos formas de almacenamiento. El grano puede guardarse después de la primera trilla en espiguillas enteras. O bien se puede realizar la segunda trilla y el grano entonces se almacena ya completamente procesado. El ensilaje de las espiguillas presenta una doble ventaja. La trilla es laboriosa. Almacenando las espiguillas y trillándolas a medida que se va necesitando grano, se reparte la tarea de trillado a lo largo del



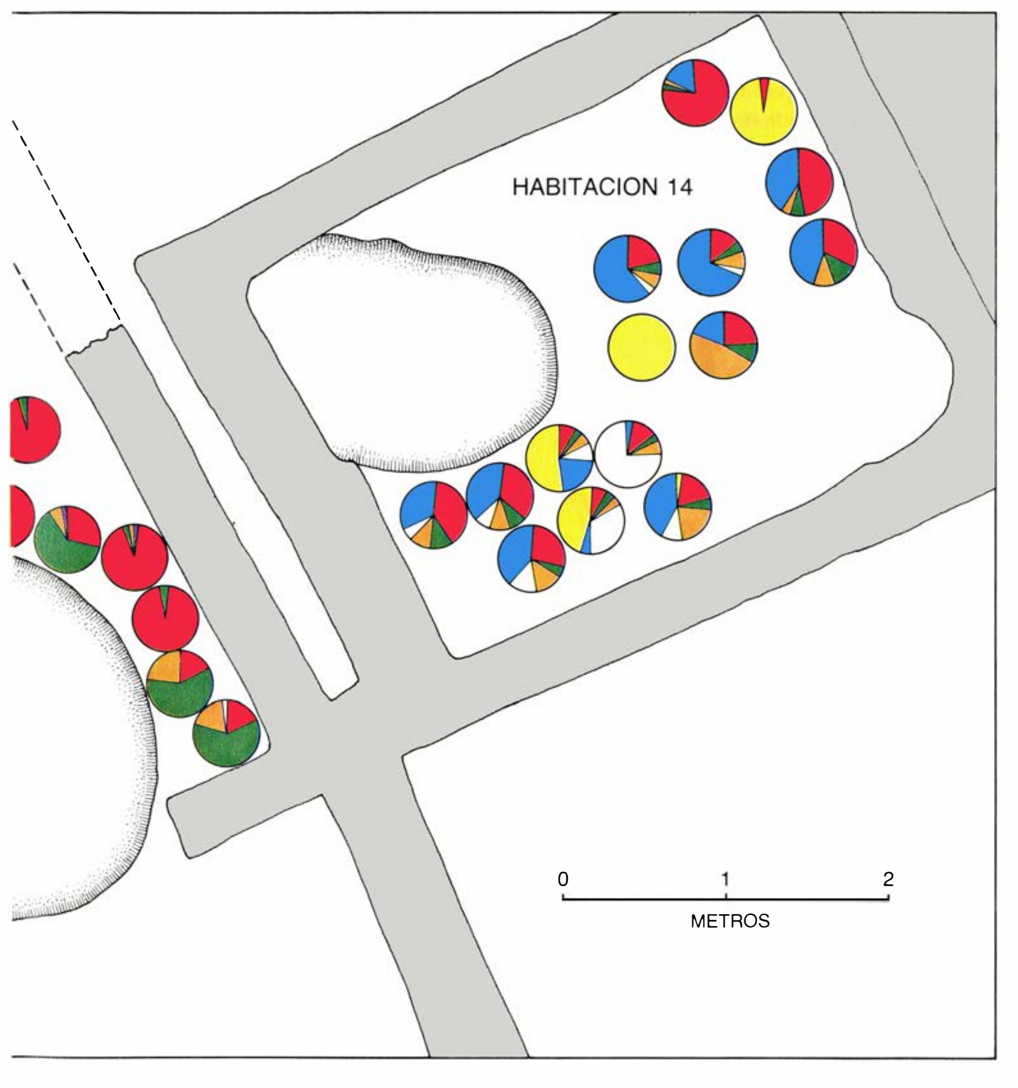

**5. GRACIAS A LAS MUESTRAS DE GRANO de los suelos de los almacenes, los arqueólogos han recabado una valiosa información sobre el ensilaje de excedentes en el yacimiento de Assiros. Se han identificado semillas de siete especies, por lo menos. Entre ellas hay cuatro tipos de trigo: esprilla, escanda, espelta y candeal. Había también otros dos cereales –cebada de seis carreras y mijo– y arvejas, afines éstas a guisantes y lentejas. Cada círculo indica el lugar**

año. Más aún, las espiguillas ofrecen mayor resistencia a los ataques de insectos y hongos que el grano. Por lo que se ve, los habitantes de Assiros aprovecharon estas ventajas y almacenaron parte de su trigo de gluma en espiguillas. Encontramos muestras de trigo de gluma con espiguillas enteras y también muestras con altas proporciones de paja.

Los muestreos exhaustivos de restos de plantas, lo hemos visto en lo hasta ahora descrito, nos han revelado el número y el tipo de especies cultivadas en Assiros; nos han enseñado la forma en que los campesinos procesaron y almacenaron sus cosechas. ¿Pueden utilizarse técnicas similares para abordar otras cuestiones menos generales, verbigracia, cuántos recipientes de cada producto había en un almacén del complejo? La respuesta es afirmativa. Al menos pueden proporcionar un cálculo del número mínimo de recipientes de almacenamiento de cada tipo de grano en un área determinada del complejo. Por ejemplo, si en un rincón de una habitación se encuentra que dos concentraciones de cebada están separadas por una concentración de esprilla, se deduce que debió haber, por lo menos, tres contenedores en aquel rincón: dos de cebada y uno de esprilla.

La razón por la cual este análisis indica solamente el número mínimo de recipientes de almacenaje es obvia: la concentración, por ejemplo de esprilla, puede contener una mezcla de granos de dos recipientes de esprilla que estaban juntos cuando el incendio los rompió. Ahora bien, combinando la estimación mínima con el recuento de hoyos de *pithoi,* podemos aproximarnos al número real de contenedores de cada tipo de grano. El ejemplo más claro de este método lo representa la habitación 9, donde se hallaron siete concentraciones separadas en la parte conservada del suelo: dos concentraciones de esprilla, dos de escanda con espelta, dos de mijo y una de arveja. Como se indicó más arriba, la habitación 9 albergaba, por lo menos, siete pozos para *pithoi.* Parece razonable concluir que hubo al menos siete recipientes de almacenaje en el área conservada de esa estancia y que la relación de concentraciones revela el contenido de los siete recipientes. Parejamente, el suelo de la habitación 14 reveló ocho concentraciones de grano y lo que bien pudieran ser diez pozos de *pithoi.* La distribución de los contenedores aquí no ofrece la nitidez que en la habitación 9, pues una o varias concentraciones de grano pudieran provenir de distintos recipientes. Sin embargo, el sistema de muestreo, cuando se combina con un examen de los restos no botánicos, permite describir de una forma plausible el ensilaje habilitado en cada habitación.

El análisis del sistema de almacenamiento en las habitaciones revela un hecho sorprendente: cada una de las tres guardaría un elenco diferente de grano. La lista de concentraciones dada en el párrafo anterior para la habitación 9 excluye el trigo candeal (o el duro), pues no se hallaron allí granos de esa especie. En cambio, trigo candeal (o duro) sólo se almacenó en la habitación 14. Por otro lado, la arveja se guardó exclusivamente en la habitación 9. El mijo está ausente de la habitación 12. La desigual distribución de productos es bastante esclarecedora y ayuda a entender cómo funcionaron los almacenes. Aunque hemos de prestar atención a otros factores antes de poder desentrañar el significado de ese sistema. El primer paso consiste en advertir que la capacidad del complejo de almacenaje es mucho mayor que la capacidad necesaria para una casa. Así las cosas, o bien cada almacén perteneció a una familia diferente o bien el complejo constituía un área de almacenamiento comunal.




**de una muestra y ofrece las proporciones de las siete especies recogidas allí. El hecho de que en algunas muestras domine nítidamente una especie da pie a afirmar que el grano no debió sufrir graves alteraciones desde el 1350 a.C. El muestreo puede revelar a los intérpretes la pauta seguida en el almacenamiento.**

Hay una buena razón para pensar que los almacenes no pertenecían a unidades familiares. La lista de especies de cada almacén excluye, al menos, una planta con propiedades altamente deseables. Por ejemplo, la cebada, siembra hoy habitual en las islas griegas como un cereal de bajo riesgo, está ausente en las habitaciones 9 y 14. El mijo (ausente en la habitación 12) se siembra en un período bajo en las faenas agrícolas. La arveja (que falta en las habitaciones 12 y 14) es bastante útil

en sistemas de rotación de cultivos. Por tanto, una familia que cosechase las especies representadas en cualquier habitación habría invertido un considerable esfuerzo perdiendo en cambio la ventaja que suponen las propiedades complementarias de toda la gama de cultivos. Más todavía, las familias que hubiesen usado las habitaciones 12 y 14 se alimentarían con una dieta bastante alejada del óptimo. La arveja, que falta en aquellas habitaciones, es un excelente complemento alimenticio de los cereales.

Supóngase, por otro lado, que los tres almacenes formaron parte de una sola unidad. Si tal hubiera ocurrido, el grupo que controlaba el complejo habría estado explotando la gama entera de cultivos disponibles. La idea de que las habitaciones constituían una unidad se apoya en el registro arquitectónico. Los pasillos que conectan los almacenes, a diferencia de la mayoría de los callejones del poblado, no están cubiertos de grava. La ausencia de ese manto sugiere que los pasillos estuvieron originariamente bajo la techumbre que uniría las habitaciones en un solo edificio. Aun así, ¿quién habría controlado esa estructura? Si atendemos al tamaño y capacidad de los almacenes, es lógico pensar que la comunidad globalmente o una parte importante de la misma poseería los granos almacenados. De acuerdo con todos los datos, lo que hemos encontrado en Assiros constituiría un auténtico fenómeno colectivo que debería haber exigido un buen grado de cooperación entre sus habitantes.

No hay razón para suponer que los productos almacenados en las habitaciones 9, 12 y 14 agotaran necesariamente las instalaciones de ensilaje en Assiros. Cada familia almacenaría quizás alimento en sus propias casas. Más que proporcionar la capacidad de almacenaje total de la comunidad, el complejo que hemos hallado fue, probablemente, un área donde se guardaba una reserva de alimento como una especie de póliza de seguro colectiva contra malas cosechas o pérdidas totales de las mismas. Un excedente de este tipo podría haberse acumulado como obligaciones en especie de la comunidad y el producto almacenado podría haber revertido a los miembros de la comunidad por los servicios prestados en la reconstrucción del muro circundante u otras tareas comunales. La naturaleza exacta de los mecanismos sociales que acompañaron a tal almacenaje colectivo todavía no pueden reconstruirse. A pesar de ello, la impresión de cooperación es viva y ayuda a trazar el cuadro de la vida del poblado en la Edad del Bronce.

La capacidad para una actividad coordinada podría aupar a Assiros por encima del nivel de simple aldea agrícola. Sin embargo, más allá de esta observación no es fácil especificar la estructura económica y social de la comunidad establecida sobre el montículo. Un modelo prometedor para comprender el asentamiento en su relación con la población de los alrededores es el de un sistema redistributivo. Es sabido que los sistemas redistributivos han existido en las primitivas sociedades del Próximo Oriente, como Sumer y Ugarit, durante la Edad del Bronce. La red de relaciones tejidas alrededor de los palacios minoicos y micénicos también se ha interpretado como un sistema redistributivo.

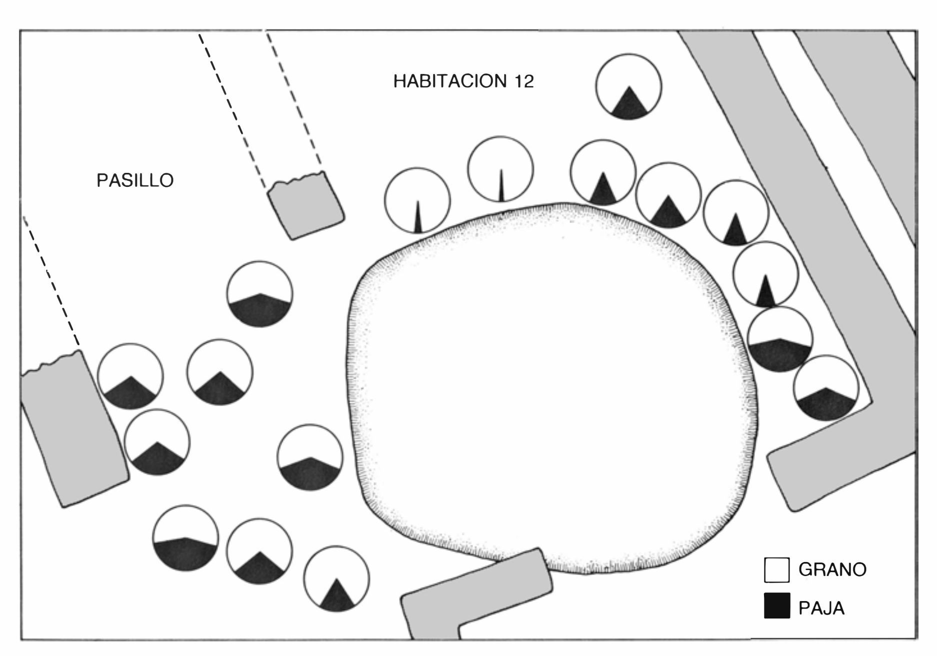
En un sistema redistributivo, las autoridades del centro, que puede ser un palacio, un templo o una ciudad, tienen el derecho de contar con ciertos bienes y servicios de la población local. En contrapartida se espera de las autoridades que proporcionen beneficios materiales, como alimento o materias primas, y beneficios menos tangibles como protección militar o la observación del ritual. Un rasgo crucial de tales sistemas de obligación recíproca, documentado en tabletas de arcilla halladas en templos del Próximo Oriente y en Cnossos y Pilos, es que no incluyen moneda o mercados. El cumplimiento de cada parte se rige por costumbres que hacen posible que bienes y servicios circulen libremente sin intercambios monetarios o trueque.

Assiros Toumba fue una localidad mucho más sencilla que los grandes palacios de la Edad del Bronce. Quizá le cayera bien el apelativo "protopalacial": el tipo de asentamiento que se habría convertido, con circunstancias favorables, en un centro redistributivo en perfecto desarrollo. Si andamos en lo cierto, nos aventuraríamos a hacer algunas predicciones sobre lo que encontramos en el curso de nuestro ulterior trabajo. Cualquier centro redistributivo, no importa cuán pequeño o primitivo sea, debe tener una red de conexiones con los asentamientos circundantes. Para descubrir qué asentamientos estuvieron en la órbita de Assiros Toumba tenemos en proyecto estudiar el área circundante.

Además, el centro sostiene, por lo general, una fracción de la población de un sistema redistributivo. Si Assiros Toumba era un sistema tal, quizás sólo viviera en el propio asentamiento una pequeña proporción de la población necesaria para construir y reconstruir el



**6. TRIGOS DE GLUMA, llamados así por las glumas o cascabillos que rodean los granos. Los cultivos de Assiros incluían tres especies: esprilla, escanda y espelta. La figura ilustra la escanda. A la izquierda se esquematiza la espiga o mazorca. Cada espiguilla, que contiene un par de granos, está unida al raquis, la parte del tallo central que está dentro de la mazorca. Al procesar los trigos de gluma, se trilla primero para romper la espiga en espiguillas (*centro*). Una segunda trilla separa los granos de la paja menuda (*derecha*). La paja menuda está constituida por ligeros fragmentos y la parte baja y dura de las glumas.**




**7. RELACION PAJA-GRANO en las muestras de trigo de gluma. De esa razón podemos inferir cómo se ensilaban las cosechas. El trigo de gluma puede almacenarse en espiguillas enteras después de una trilla o en forma de grano completamente procesado tras dos trillas. Si se almacena después de la primera trilla, la mayor parte de la paja menuda todavía rodea al grano. La alta proporción de paja en muestras obtenidas de la habitación 12 indican que al menos parte del trigo de gluma de Assiros se ensiló de esa forma.**

muro. El resto de la población se alojaría en los asentamientos circundantes y habría contribuido con su esfuerzo en el levantamiento del muro defensivo a cambio de protección, comida u otras contrapartidas. Es posible que la población del propio Assiros Toumba se cifrase únicamente entre 50 y 100 individuos. Serán necesarias más excavaciones dentro del muro defensivo para saber si éste fue el caso.

Los sistemas redistributivos constituyen uno de los puntos de mayor interés teórico por parte de arqueólogos y antropólogos. En este sentido, el trabajo de Assiros Toumba podría prestar, en última instancia, una útil contribución a ese capítulo. Cuestión importante es desentrañar el origen de los sistemas redistributivos. Si bien se trata de un tema abierto, existe una hipótesis según la cual estos sistemas protegen a las comunidades agrícolas contra el hambre, al instar el ensilaje del excedente. Una ratificación impresionista de esta hipótesis nos la ofrece el relato bíblico de José. Al seguir su sueño y almacenar el excedente de siete años de «vacas gordas» para prevenir siete años de «vacas flacas», José estaba cumpliendo las obligaciones del centro en un sistema redistributivo.

Si consideramos el complejo de almacenes como una práctica complementaria, más que sustitutiva, del almacenaje en las casas, tendría sentido la idea de que la redistribución vendría a paliar las oscilaciones de la economía agrícola. El ensilaje cumpliría la misión de reserva y no se trataría de un mero sistema centralizado de almacenaje colectivo para uso cotidiano. En este sentido, el trabajo acometido en Assiros Toumba podría iluminar algunas cuestiones teóricas de importancia. Esa posibilidad queda para el futuro.

De momento podemos afirmar que los trabajos en el montículo han superado ya con creces lo que esperábamos averiguar. Empezamos nuestra excavación sin apenas conocer qué restos encerraba Assiros. Parecía razonable pensar en huellas de una típica aldea de la Edad del Bronce con unas cuarenta o cincuenta casas. Creíamos que esas viviendas serían unidades independientes, que contaban con su propio trabajo y sin los mecanismos necesarios para empujarlas hacia una comunidad mayor. En cambio, descubrimos una comunidad que ponía de manifiesto mecanismos complejos de cooperación social, incluido quizás un sistema redistributivo que subsumía a la vez trabajo y excedente agrícola. Bajo esta perspectiva Assiros reviste mucho mayor interés, porque se asemeja tanto a las grandes ciudades palaciales del sur de Grecia como a una simple aldea agrícola. En los años venideros nuestro conocimiento del origen de las economías palaciales de la Edad del Bronce irá consolidándose y, en esa tarea, Assiros aportará, sin la menor duda, una gran contribución.

# El efecto Hall cuántico

*La cuantización de la resistencia de Hall que, a bajas temperaturas y en campos magnéticos intensos, ofrecen determinados componentes de los semiconductores guarda relación con diversas constantes fundamentales de la naturaleza*

Klaus von Klitzing

No es cosa habitual que se conozca con precisión el instante del nacimiento de un efecto físico; ése es, sin embargo, el caso del efecto Hall cuántico. Se produjo durante la noche del 4 al 5 de febrero de 1980, hacia las 2 de la madrugada, en el laboratorio de electroimanes de campo intenso de Grenoble: de repente se cayó en la cuenta de que los resultados experimentales debían interpretarse de un modo enteramente nuevo.

Como en las semanas y meses precedentes, el autor y su colaborador Thomas Englert investigaban la influencia que ejerce un campo magnético sobre las propiedades electrónicas de transistores *MOS* de efecto de campo (semiconductor de metal-óxido), el componente principal de la moderna microelectrónica [*véase la figura 1*]. El laboratorio de electroimanes de campo intenso del Instituto Max-Planck de Física del Estado Sólido, que se utiliza conjuntamente con el Centro Nacional francés de Investigación Científica (CNRS), resulta particularmente adecuado para esos experimentos, pues se generan allí campos magnéticos de más de 20 tesla, es decir, de unas 400.000 veces el campo magnético terrestre.

En los campos magnéticos de tamaña intensidad, y a bajas temperaturas, se modifican drásticamente las propiedades electrónicas de los transistores. Aparecen, por ejemplo, oscilaciones de la resistencia eléctrica en función del campo magnético. Tales mediciones no constituían novedad, pues desde hacía más de 20 años se venían desarrollando experimentos de ese tipo en diversos laboratorios, principalmente en los Estados Unidos, Japón, Inglaterra y la República Federal de Alemania.

El primer objetivo de esas investigaciones era recabar nueva información, a partir del análisis de las mediciones, sobre complejos detalles microscópicos de los elementos constitutivos de los semiconductores. Cuestión especialmente importante en ese contexto es determinar qué perturbaciones y mecanismos entorpecen en el material semiconductor el movimiento de los electrones, y con ello la velocidad de transmisión de información.

Con sorpresa advertimos aquella madrugada que ciertas propiedades eléctricas de los elementos constitutivos del semiconductor sometido a un campo magnético intenso *no* dependían de la morfología del cuerpo sólido. Tanto si las mediciones se efectuaban con muestras proporcionadas por Gerhard Dorda, de los laboratorios Siemens, como con las de Mike Pepper, de la compañía Plessey, se obtenía siempre una resistencia eléctrica de alrededor de 6453,2 ohm, valor que debía haber fijado la naturaleza. A lo que parecía, las características microscópicas distintivas de los elementos constitutivos del semiconductor no ejercían influencia alguna sobre el resultado.

El resultado experimental se detalla en la figura 2. Una corriente atraviesa una "caja negra" que caracteriza la muestra. Como en el caso de la resistencia eléctrica común, se mide la tensión, que crece proporcionalmente con la intensidad de la corriente. La constante de proporcionalidad establece el valor de la resistencia; en general, eligiendo adecuadamente el material por el que fluye la corriente y su geometría, la resistencia puede tomar cualquier valor.

Existen, sin embargo, ciertas muestras (componentes especiales de semiconductores) cuyos valores de resistencia aparecen cuantizados. En ese sentido, suele decirse que presentan una resistencia de Hall cuantizada o, en términos generales, que muestran efecto Hall cuántico. La resistencia medida no toma ya cualquier valor arbitrario; antes bien, ciertos valores resultan especialmente estables.

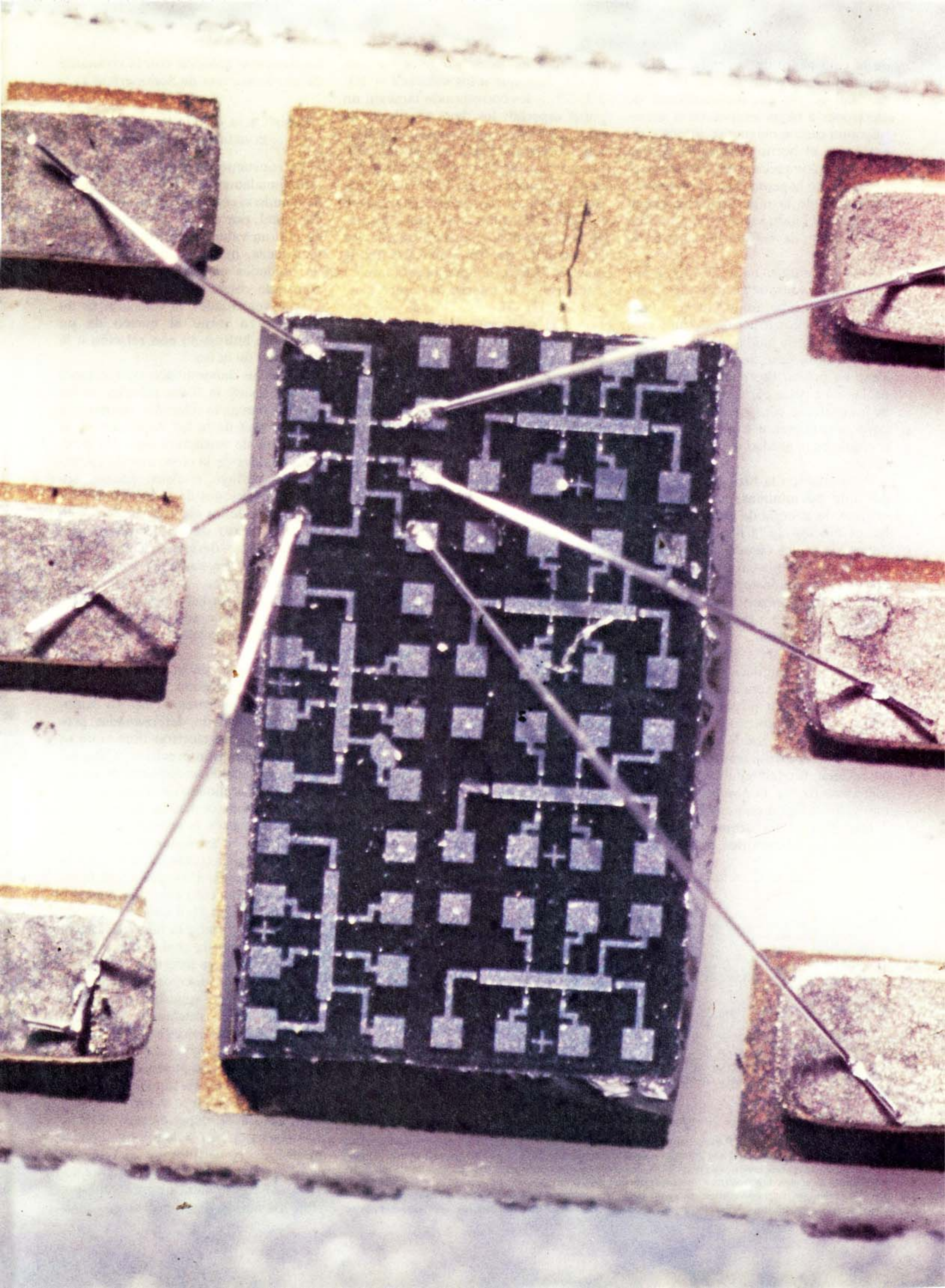
Hasta la fecha, todos nuestros estudios revelan que, con una precisión de hasta la millonésima, esos valores cuantizados dependen únicamente de constantes naturales; por consiguiente, en cualquier lugar del mundo pueden reproducirse con igual precisión. Parece milagrosa la existencia de tal "resistencia natural". Para el éxito de nuestra medición resulta decisivo que el valor de la resistencia, aproximadamente 6453,2 ohm, amén de poderse generar con especial facilidad, se encuentre precisamente en una región accesible a las técnicas de medición.

Los métodos de precisión con que se cuenta hoy para calibrar resistencias resultan idóneos en la gama de valores comprendidos entre 1 ohm [*véase la figura 3*] y 10.000 ohm. Para medir resistencias menores debe trabajarse con corrientes muy grandes, mientras que para resistencias mayores los elementos aislantes pueden originar corrientes de pérdida.

## Campos magnéticos intensos y temperaturas bajas

Si queremos obtener la resistencia natural deben, por supuesto, determinarse con precisión las características

**1. SE DESCUBRIO EL EFECTO HALL cuántico midiendo transistores de silicio de efecto de campo. El efecto Hall clásico se remonta a 1880; ese año el físico norteamericano Edwin Herbert Hall lo describió en un conductor recorrido por una corriente y sometido al influjo de un campo magnético perpendicular al mismo. El montaje en el que se registra el efecto Hall cuántico es, en principio, semejante, salvo en que el conductor se ha sustituido por un transistor de efecto de campo. En un campo magnético suficientemente intenso, y a muy bajas temperaturas, la tensión y la resistencia de Hall muestran un claro escalonamiento. Sobre la lámina de silicio que aparece en la figura se reconocen siete transistores, de 0,4 milímetros de longitud cada uno; uno se muestra aquí conectado a conductores de medición. Se fabricaron en los laboratorios Siemens para emplearlos en investigación; el tamaño de los correspondientes componentes de uso habitual en microelectrónica unas cien veces más pequeños.**



de la caja negra de la figura 2. El especialista suele referirse a la resistencia de Hall de un gas bidimensional de electrones a bajas temperaturas sometido a un campo magnético intenso; sin embargo, el hecho destacable es que tanto el valor preciso del campo magnético como la geometría, el material o la calidad de la muestra desempeñan un papel del todo secundario a la hora de averiguar la resistencia cuantizada de Hall.

Dos constantes naturales fijan por completo el valor de la resistencia; ambas revisten una importancia fundamental en la construcción del universo físico a partir de unidades menores. En este sentido, casi siempre suele pensarse en primer lugar en los átomos o en las partículas elementales, a partir de las cuales se construyen aquéllos, en cuyo caso la menor carga eléctrica observable es la unidad de carga elemental, *e*.

Pero también la luz está constituida por unidades mínimas, por paquetes de energía; la energía de tales cuantos de luz es función de la longitud de onda de la luz y del cuanto de acción de Planck, *h*. Entre la carga elemental *e* y el cuanto de acción de Planck puede establecerse una razón cuyas dimensiones corresponden a la de la resistencia: $h/e^2 = 25812{,}8$ ohm.

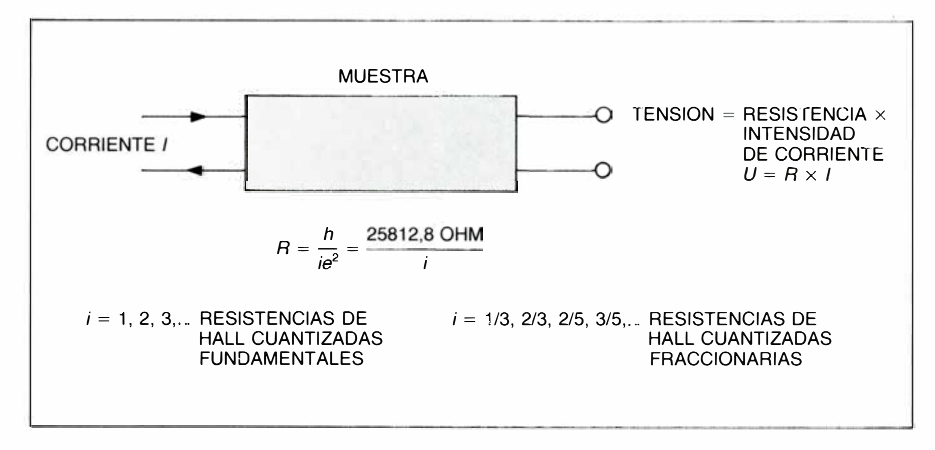
El cociente $h/4e^2 \approx 6453{,}2$ ohm proporciona el valor de la resistencia de Hall cuantizada, sometido hasta la fecha a un examen particularmente intenso. En principio, sin embargo, pueden reproducirse todos los valores de $h/ie^2$; en el caso de la resistencia de Hall cuantizada fundamental, *i* es un número entero. En la práctica se examinan valores de *i* menores que doce.

Dan Tsui, Horst Störmer y Art Gossard, de los laboratorios Bell, han demostrado, en investigaciones no concluidas aún, que a los valores $i = 1/3$, $2/3$, $2/5$,... les corresponde también un papel especial; los denominadores de esas fracciones son siempre impares. Las teorías que hasta hoy se han propuesto para explicar ese efecto, denominado resistencia de Hall cuantizada fraccionaria, parten de una idea de Bob Laughlin. Resultan relativamente complejas, pues debe considerarse a la vez la totalidad de los electrones de la muestra. Especialmente importante en este caso es la interacción entre electrones.

Por el contrario, la resistencia de Hall cuantizada fundamental puede discutirse en un marco bastante simple donde se consideran por separado los electrones. Antes de abordar el efecto Hall cuántico, sin embargo, quiero detenerme en el significado práctico del fenómeno cuántico.

En la hipótesis de que exista una resistencia natural de valor fundamental $h/e^2$, y de que, además, pueda observarse experimentalmente, cabría determinar esa constante fundamental con medios sencillos [*véase la figura 4*]. A tal efecto se compara la caída de tensión $U_1$ en una resistencia conocida, $R_1$, con la caída de tensión $U_2$ en la resistencia natural. Por ambas resistencias pasaría la misma intensidad de corriente. La medición arrojaría directamente el valor de $h/e^2$.

En diversas universidades de los Estados Unidos, Inglaterra y Alemania se realiza hoy ese experimento en las clases de prácticas de física. Lo notable del valor de $h/e^2$ es que, hasta un factor numérico conocido con exactitud (que, en esencia, contiene el valor, determinado internacionalmente, de la velocidad de la luz, a saber, $c = 299792458$ metros por segundo), esa constante fundamental coincide con la constante de estructura fina de Sommerfeld.




**2. EN LA RESISTENCIA DE HALL cuantizada, la caída de tensión, *U*, es, como en el caso de una resistencia eléctrica normal, proporcional a la intensidad de la corriente de ensayo, *I*, si bien la constante de proporcionalidad, *R*, toma únicamente ciertos valores. Estos dependen exclusivamente de las constantes naturales *h* (el cuanto de acción de Planck) y *e* (la carga eléctrica elemental). La cuantización de la resistencia sólo se observa en campos magnéticos intensos, a muy bajas temperaturas y en muestras en las cuales la corriente atraviesa una delgadísima lámina. La "caja negra" simboliza una muestra de esas características.**

## Resistencia de Hall cuantizada y constantes naturales

La constante de estructura fina vale aproximadamente 1/137. Considerada a menudo la constante natural más fundamental, por carecer de dimensiones, es sólo un valor numérico dado por la naturaleza, independiente del sistema de unidades elegido. El valor de la constante corresponde al cociente entre dos velocidades: la velocidad de un electrón en torno al núcleo de un átomo de hidrógeno con relación a la velocidad de la luz.

Todas las denominadas correcciones relativistas de la física atómica en las que intervenga la velocidad respecto de la velocidad de la luz dependen de la constante de estructura fina. En puridad, el valor de la constante de estructura fina influye, en física atómica, sobre las posiciones de todos los niveles de energía de los electrones. Pero sólo en el contexto de la electrodinámica cuántica, que describe la interacción entre partículas cargadas y radiación electromagnética, podían efectuarse predicciones verdaderamente precisas. Se da aquí por supuesto el conocimiento de la constante de estructura fina.

En el marco de la electrodinámica cuántica, y haciendo uso de ordenadores muy potentes, Toichiro Kinoshita ha calculado determinadas propiedades de un electrón libre, que se midieron con gran precisión. Cabe esperar que cuando se disponga de un valor bien determinado de la constante de estructura fina (deducido, por ejemplo, de mediciones de la resistencia de Hall cuantizada) se adviertan posibles desviaciones de la teoría de la electrodinámica cuántica. Hay quien señala incluso que, conocido ese valor exacto a partir del efecto Hall cuántico, podría someterse a prueba la hipótesis, aceptada hoy, según la cual el electrón carece de estructura interna.

En cualquier caso, todo nuevo método de medida aplicable a la determinación de constantes naturales presta excelentes servicios a la hora de validar teorías y experimentos. A menudo, un nuevo método de determinación de constantes fundamentales ha puesto en cuestión la valoración e interpretación de experimentos anteriores. Haciendo uso de la totalidad de las mediciones y cálculos de que se dispone, casi cada diez años se intenta determinar con la menor imprecisión posible los valores de las principales cons-

tantes naturales. Los resultados de la estimación más reciente se harán públicos este mismo año.

¿Dónde se hallan, pues, los límites de la precisión de un experimento? ¿Con qué exactitud puede determinarse la constante de estructura fina a partir de la resistencia de Hall cuantizada, aun suponiendo que el valor de la resistencia sea $h/e^2$? El problema equivale a preguntar con qué precisión puede medirse el valor de una resistencia eléctrica. El montaje de medición muestra en principio [*véase la figura 4*] que deben conocerse, por una parte, la relación $U_2/U_1$ entre dos tensiones y, por otra, el valor exacto de la resistencia $R_1$ con la que se establece la comparación.

Resulta fácil la medición de tensiones relativas con incertidumbres inferiores a una parte en cincuenta millones, sobre todo si las tensiones a comparar son similares (esa misma incertidumbre en un cálculo de la población española significaría que el error sería inferior a un ciudadano). El valor exacto de un patrón de resistencia resulta hoy, en comparación, al menos diez veces más impreciso. La dificultad radica, por tanto, en obtener una resistencia cuyo valor se conozca con exactitud.

Sobre el papel, todas las unidades físicas se determinan de manera unívoca y sin fallo a partir de la fijación de unidades básicas y unidades derivadas, las cuales se obtienen a partir de las leyes físicas; sin embargo, en la práctica experimental se producen siempre imprecisiones de medida.

En el caso de la unidad básica de electricidad (la unidad de intensidad de corriente, o ampère) la imprecisión no es pequeña debido a la dificultad que entraña llevar a la práctica la definición del ampère: "un ampère es la intensidad de una corriente invariable que fluye a lo largo de la doble conducción constituida por dos conductores infinitamente largos, de secciones despreciablemente pequeñas, rectilíneos y paralelos, y que se encuentran a una distancia de un metro, cuando la fuerza producida por la corriente entre ambos conductores en el vacío es de $2 \times 10^{-7}$ newton por cada metro de longitud de la doble conducción".

Ni un solo laboratorio del mundo está hoy en condiciones de producir una corriente de un ampère con una precisión de un microampère. Incluso los científicos del Laboratorio Nacional de Física de Teddington, cerca de Londres, especializados en medir con una balanza de corriente la fuerza que se genera entre dos alambres al paso de la electricidad, y obtener así información directa sobre la intensidad de la misma, parten del hecho de que no puede generarse una corriente de 100 ampère con precisión inferior al miliampère.



**3. UNIDAD DE RESISTENCIA ELECTRICA: ohm. Suelen disponer de ella las oficinas públicas de pesos y medidas en forma de resistencias alámbricas. La figura muestra una resistencia de precisión de un ohm; su valor real se altera con el paso del tiempo, pues el material del alambre no resulta estable a largo plazo.**

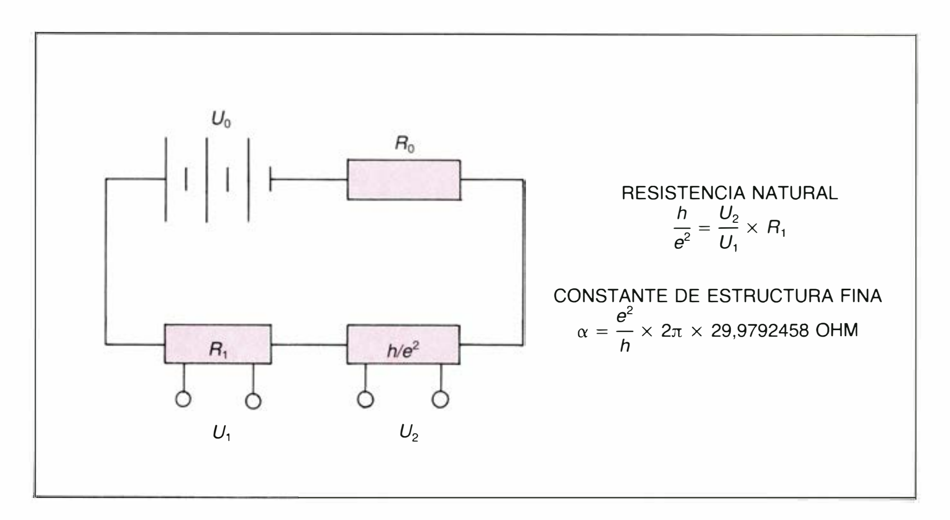
Afortunadamente se consiguen resistencias de un ohm con precisiones 100 veces mayores que la de la unidad básica de intensidad, puesto que no sólo puede crearse el ohm combinando corriente y tensión (en cuyo caso la imprecisión sería cuando menos tan grande como la del ampère), sino también mediante la corriente de intercambio-resistencia de un condensador.

En tal caso deben conocerse la frecuencia de la corriente de intercambio y la geometría del condensador, de modo que, por este procedimiento, bastan las unidades básicas de tiempo y longitud, que sí se conocen con suficiente precisión. Unos cinco laboratorios de todo el mundo se encuentran hoy en condiciones de producir el ohm con una precisión de una parte en diez millones.

## Determinación precisa del valor de la resistencia

Las resistencias normalizadas de alambre que suelen conservar como patrón muchos países [*véase la figura 3*] son imprecisas, pues el material del alambre no es estable a largo plazo y el valor de la resistencia varía, pues, con el tiempo. Cualquier oficina de pesos y medidas estaría encantada si pudiese establecer en cualquier momento el ohm con una precisión superior a una millonésima.

En las mediciones para determinar con precisión el valor de la resistencia de Hall cuantizada participaron fundamentalmente el Instituto Federal Físico-Técnico de Braunschweig (Repú-




**4. DETERMINACION DE LA CONSTANTE FUNDAMENTAL $h/e^2$. Para ello deben efectuarse dos mediciones de tensión, $U_1$ y $U_2$. Debe conocerse también el valor de la resistencia $R_1$, mientras que la primera resistencia $R_0$ y la fuente de tensión $U_0$ se eligen de tal manera que fluya una corriente de menos de una diezmilésima de ampère. La constante de estructura fina, $\alpha$, es idéntica a la cantidad que se mide, $e^2/h$, hasta un factor invariable que es función del valor de la velocidad de la luz, fijado éste internacionalmente.**

blica Federal de Alemania), la Oficina Federal de Pesos y Medidas de los Estados Unidos y el Laboratorio Electrotécnico del Japón. Se observó repetidamente que la precisión y la estabilidad de la resistencia patrón, con la que se comparaba la resistencia de Hall cuantizada, influía de manera decisiva en la precisión total de la medición.

Probablemente nunca pueda determinarse la constante de estructura fina con ayuda del efecto Hall cuántico con una precisión mayor que la de la resistencia con la cual se compara. Sin embargo, tras comprobarse que la resistencia de Hall cuantizada resulta más estable y de más fácil reproducción que cualquier resistencia alámbrica, se nos ofrece otra aplicación del efecto Hall cuántico: cabe plantearse que en todo el mundo se reconozca como patrón la resistencia natural $h/e^2$, pues su valor es estable en el tiempo e independiente del lugar.

Ya hoy puede medirse la variación temporal del valor de resistencias de precisión con ayuda de la resistencia de Hall cuantizada. Como valor numérico se ofrece el de $h/e^2$, calculado a partir del que se recomiende internacionalmente para la constante de estructura fina. Una propuesta al efecto será objeto de discusión este año en la próxima sesión de las asociaciones competentes en unidades eléctricas. En laboratorios estatales de al menos 16 países se investiga actualmente las posibilidades de utilización de la resistencia de Hall como patrón normalizado, y hay buenas perspectivas de que en el transcurso de este mismo año se posean datos suficientes que justifiquen tal empleo.

Al trazar la evolución habida en la fijación de unidades físicas se reconoce la tendencia a relacionar tales unidades con magnitudes naturales invariables. Las medidas corporales, como el pie o el codo, no satisfacían esas condiciones, dado que tales magnitudes, incluso en promedio estadístico, dependen del tiempo y del lugar.

Durante largo tiempo el hombre creyó que nuestro sistema planetario constituía una magnitud natural adecuada: constante el tiempo de rotación de la tierra alrededor del sol y fija la circunferencia terrestre; de ellos se dedujeron, por tanto, las unidades de tiempo y de longitud. Hoy sabemos que esas cantidades no son estables, mientras que admitimos que las propiedades atómicas y las constantes naturales no varían con el tiempo y son las mismas en cualquier lugar del universo.

Las unidades físicas que se apoyan en datos atómicos ofrecen, así, la gran ventaja de que cualquiera, en cualquier instante, puede trabajar con la misma medida física sin estar supeditado al intercambio de aparatos de medida. Por esa razón constituyen hoy las frecuencias atómicas (el reloj de cesio) el fundamento de la unidad de tiempo y, la velocidad de la luz, como constante natural inmutable, junto con la unidad de tiempo, la unidad de longitud.

La masa del átomo se convertiría en la magnitud ideal para fijar la unidad de masa; desgraciadamente, no se puede determinar aún el número de átomos de un cuerpo macroscópico con la suficiente exactitud. Pese a lo cual se trabaja con ahínco para determinar con máxima precisión la distancia a la que se encuentran los átomos en los monocristales de silicio; puesto que guardan una ordenación regular, podría estimarse, a partir de este dato, el número de átomos del cristal.

Sin embargo, con toda seguridad, y al menos durante los próximos diez años, el venerable kilogramo de la Oficina Internacional de Pesos y Medidas (BIPM) de Sèvres, en los alrededores de París, seguirá constituyendo el patrón de referencia para la masa, dado que la elaboración de unidades de kilogramo por comparación de masas mediante una balanza resulta mucho más fácil que la fabricación de una masa igual atendiendo al número de átomos que contenga.

En el caso de la resistencia eléctrica se observa, por el contrario, que hoy mismo cuesta menos fabricar una resistencia a partir de constantes naturales (la resistencia de Hall cuantizada) que reproducirla por otros métodos. Resulta muy probable, pues, que en el futuro se cuente por doquier, en la medición de resistencias a través de la resistencia de Hall cuantizada, con un instrumento de medida mejor que el empleado hasta el presente.

Condición imprescindible en todas estas consideraciones es que la resistencia de Hall cuantizada dependa exclusivamente de constantes fundamentales; no es admisible desviación sistemática alguna, en particular ninguna que dependa de la muestra. En condiciones ideales resulta relativamente fácil comprender, que con ayuda de determinados materiales semiconductores, pueda fabricarse una resistencia de valor $h/e^2$; en lo que sigue doy de ello, en forma simplificada, una interpretación física.

## El efecto Hall clásico

El físico norteamericano Edwin Herbert Hall publicó en 1880, es decir, exactamente 100 años antes del descubrimiento del efecto Hall cuántico, la primera descripción del comportamiento de un conductor de una corriente bajo la influencia de un campo magnético perpendicular al mismo. Su trabajo constituye la base del efecto que aquí se estudia.

La intensidad de la corriente eléctrica a través de un conductor depende de la concentración de electrones que pueden moverse y de su velocidad media [*véase la figura 5*]. La particularidad del efecto Hall es que, al aplicar un campo magnético perpendicular al conductor, se crea un campo eléctrico perpendicular al campo magnético y perpendicular también al flujo de la corriente.

Ese campo, denominado de Hall, resulta necesario para que se mantenga la

dirección inicial de la corriente. Concretamente, al estar inmersos en el campo magnético $B$, sobre los electrones actúa una fuerza perpendicular al campo magnético y perpendicular a su velocidad. Tal fuerza, denominada de Lorentz, debería, en principio, alterar la dirección de la corriente. Ahora bien, dado que ésta no puede rebasar el borde lateral del conductor, se crea un campo de Hall, que compensa la fuerza de Lorentz. Este campo es susceptible de medición en forma de una tensión de Hall, $U_H$, perpendicular al campo magnético y a la dirección de la corriente.

Se designa por resistencia de Hall el cociente entre la tensión de Hall y la intensidad de la corriente empleada. La resistencia de Hall, $R_H$, no lo es en el sentido ordinario, puesto que la intensidad de corriente y el campo eléctrico correspondiente son perpendiculares entre sí. Aumenta la intensidad del efecto Hall al hacerlo el campo magnético $B$ y al disminuir la densidad de portadores de carga, $n$, (crece la velocidad de los electrones al mantenerse constante la intensidad de corriente). En la resistencia de Hall se cumple por tanto la importante relación $R_H \sim B/n$: es proporcional al cociente entre el campo magnético y la densidad de portadores de carga.

La aplicación más importante del efecto Hall clásico es la determinación de la densidad de portadores de carga en materiales capaces de conducir la electricidad, especialmente en los semiconductores, en los que, mediante ciertos procesos técnicos (punteado), puede modificarse a voluntad la densidad de portadores de carga.

En la versión clásica no sobresale especialmente ningún valor de la resistencia de Hall. Al incrementarse el campo magnético crece la resistencia de forma monótona [*véase la figura 6*]; la pendiente de las rectas correspondientes depende de la concentración de portadores de carga.

Por el contrario, en el efecto Hall cuántico se observa una estabilización de la resistencia de Hall en los valores $h/e^2$, $h/2e^2$, $h/3e^2$, etcétera. La aparición de esas "mesetas" de Hall está vinculada al hecho de que, en la correspondiente región del campo magnético, puede producirse un flujo de corriente sin pérdida disipativa. En la práctica experimental el fenómeno se manifiesta porque desaparece la caída de tensión en la dirección del flujo de corriente. Un comportamiento similar se conoce en superconductividad, donde se observa asimismo una corriente sin pérdidas disipativas.

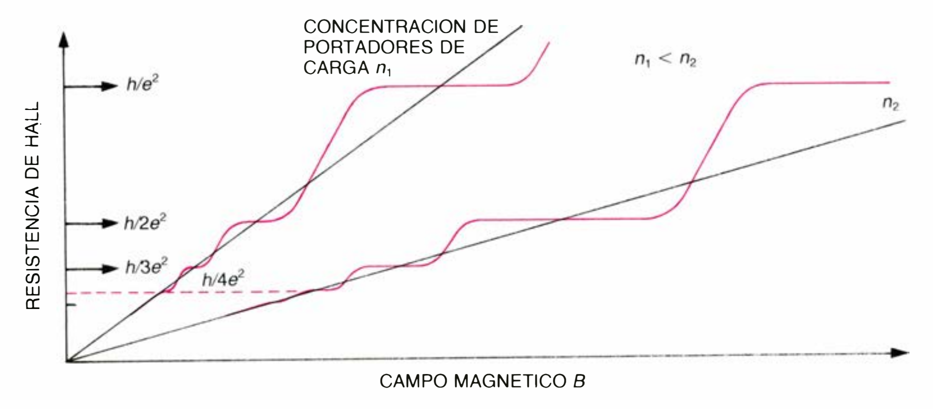
En ambos casos puede, cuando menos teóricamente, mantenerse una corriente duradera en un anillo sin aporte de energía. Sin embargo, no resulta factible llevarlo a la práctica en el caso del efecto Hall cuántico: se puede, por supuesto, doblar un conductor eléctrico hasta formar un anillo cerrado, pero resulta imposible generar un campo magnético cuya dirección siempre sea perpendicular al anillo [*véase la figura 7*].

## Modelos teóricos

A pesar de todo, suele utilizarse ese esquema para explicaciones teóricas del efecto Hall cuántico. La cuantización de la resistencia de Hall se razona a partir de proposiciones basadas en principios fundamentales (descripción mecánico-cuántica de los electrones en un campo magnético, teorías de la in-



**5. A LA INTENSIDAD DE CORRIENTE, *I*, que fluye por un conductor eléctrico le corresponde una velocidad media, *v*, de los electrones, dotados de libertad de movimiento (*a*). Si se dispone un campo magnético *B* perpendicular a la dirección de la corriente, se ejerce sobre los electrones una fuerza, *K*, perpendicular a la dirección de la corriente y perpendicular también al campo magnético (*b*). Esa fuerza, denominada de Lorentz, que se ejerce sobre los electrones del conductor eléctrico debe compensarse por medio de una tensión transversal $U_H$ (llamada de Hall), puesto que los electrones, aun en presencia del campo magnético, prosiguen su viaje en la dirección original. La resistencia de Hall, $R_H$, viene dada por el cociente entre la tensión de Hall, $U_H$, y la corriente, *I*.**




**6. RESISTENCIAS DE HALL, clásica (*líneas negras*) y cuantizada (*curvas rojas*), representadas aquí como funciones del campo magnético *B* para dos conductores distintos con diferentes concentraciones de portadores de carga, $n_1$ y $n_2$, respectivamente. Las pendientes, en ambos casos, dependen de esas concentraciones.**

variancia de aforo o "gauge"). Sin duda resultan elegantes esos métodos, pero también poco comprensibles, por lo que en este artículo daré una interpretación algo más intuitiva, basada en el efecto Hall clásico.

El supuesto más importante, común a todas las teorías del efecto Hall cuántico, es que en un conductor eléctrico inmerso en un campo magnético se produce una corriente duradera sin pérdidas disipativas. Se comprobará que ello sólo es posible para ciertos valores de la razón entre intensidad del campo magnético y densidad de portadores de carga. Dado que la resistencia de Hall también depende únicamente de esa razón, se obtienen para ella determinados valores, que caracterizan la resistencia de Hall cuantizada.

La corriente sin pérdidas del efecto Hall cuántico no debe compararse con la conductividad infinitamente buena de la superconductividad, pues las nociones de resistencia eléctrica y de conductividad carecen de significado sencillo en presencia de un campo magnético. Si se define, por ejemplo, la conductividad de una substancia como la intensidad de la corriente en la dirección del campo eléctrico aplicado, se concluye entonces que para una resistencia de Hall cuantizada (como para un aislante) la conductividad es nula.

Por consiguiente si, con un acumulador, se aplica tensión a los dos bordes de un dispositivo anular [*véase la figura 7*], cuando se satisfacen las condiciones del efecto Hall cuántico no se toma corriente de la batería. El sistema se comporta como un aislante: su conductividad es nula. Sin embargo, al contrario que en un aislante, se genera un flujo de corriente *perpendicular* al campo eléctrico.

¿Por qué desaparece toda conductividad de un conductor eléctrico? La respuesta exige analizar con precisión los procesos de transporte de portadores de carga en campos eléctricos y magnéticos. Esa constituye la base de una comprensión teórica (incompleta aún en sus detalles) de la resistencia de Hall cuantizada.

En el nivel microscópico, cuando la corriente fluye a través de un metal los electrones se aceleran en el campo eléctrico, frenándose por choque al cabo de cierto recorrido libre y acelerándose luego de nuevo. Se establece en el medio una velocidad de arrastre de los electrones y, con ello, una corriente; ésta resulta particularmente intensa cuando el tiempo que media entre dos choques se alarga. Así, cuantos menos choques se produzcan, mayor será la conductividad.

Sin embargo, en presencia de un campo magnético intenso, la situación se invierte. Deben ahora imprescindiblemente producirse choques para que se dé un transporte de electrones en la dirección del campo eléctrico. En el marco de la electrodinámica clásica la explicación es sencilla: inmersos en un campo magnético los electrones se mueven en trayectorias circulares situadas en el plano perpendicular al campo magnético; al aplicar un campo eléctrico, el centro de cada trayectoria circular se halla, en el caso ideal (esto es, en ausencia de interferencias), sobre una línea que discurre perpendicularmente al campo eléctrico [*véase la figura 8*].

El choque perturba el movimiento circular: el electrón inicia una nueva trayectoria circular en torno a un nuevo centro, desplazado en la dirección del campo eléctrico. Cuantos más choques se produzcan (con el consiguiente desplazamiento del centro del círculo) mayor será el transporte de electrones en la dirección del campo eléctrico y, por tanto, mayor la conductividad.

Ahora bien, si los centros de las trayectorias circulares no pueden desplazarse en la dirección del campo eléc-

**7. EN EL ANALISIS TEORICO de la resistencia de Hall cuantizada suele emplearse un dispositivo anular. Al aplicar una tensión de Hall $U_H$ a las aristas laterales del anillo se genera una corriente de Hall $I$; si las condiciones para el efecto Hall cuántico se satisfacen, dicha corriente fluye sin pérdida energética. Del acumulador, que rinde una tensión $U_H$, no se toma corriente alguna ($I_0 = 0$), puesto que no se precisa aprovisionamiento de corriente: el sistema de ensayo se comporta en ese sentido cual si se tratara de un aislante.**

trico, porque los electrones no encuentran "sitio libre" (estado electrónico) al que mudarse, la conductividad desaparece y el efecto Hall cuántico se torna mensurable. Por ejemplo, un electrón no puede cambiar a un estado que se encuentre ya ocupado por otro electrón (lo impide el principio de exclusión de Pauli, que dice: dos niveles electrónicos, cada uno de ellos definido por cierto número de características, denominadas números cuánticos, no pueden tener idénticas características.)

Consecuencia de ese principio es, por ejemplo, que los electrones de un átomo no pueden moverse todos sobre la misma trayectoria en torno al núcleo atómico: deben ocupar capas distintas. Restringiéndonos a un modelo atómico sencillo, los electrones que se hallan en una de esas capas tienen todos la misma energía; los átomos de los gases nobles, así helio o neón, son particularmente estables porque todos los estados de sus respectivas capas externas se encuentran completamente ocupados por electrones.

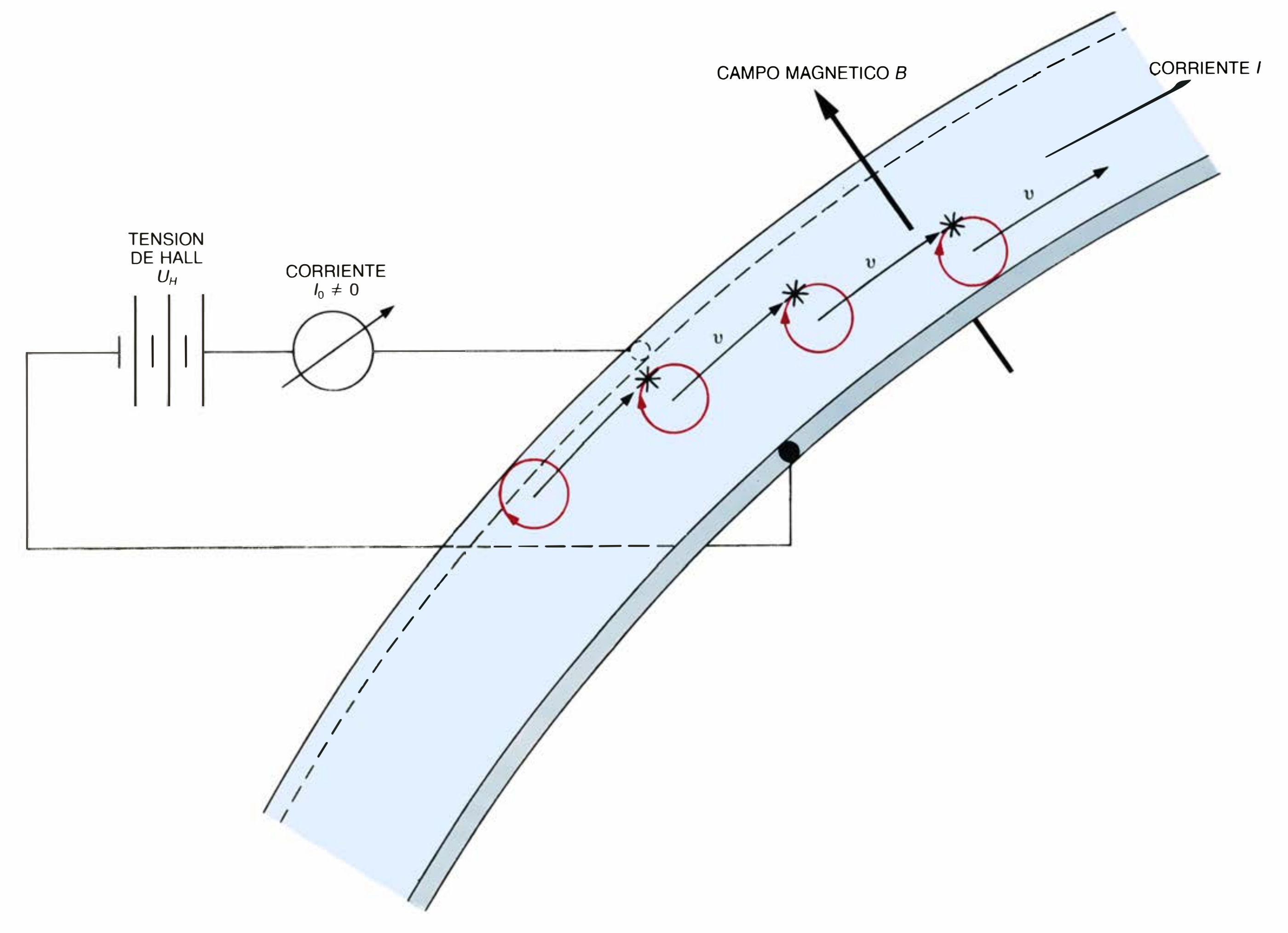
Análogamente, las peculiaridades del efecto Hall cuántico guardan relación con dos datos: un nivel energético se encuentra totalmente ocupado por electrones y entre los estados ocupados y los no ocupados media un vacío energético relativamente grande. Al contrario que en el caso del átomo aislado, el cálculo del espectro energético de los electrones de un cuerpo sólido real (metal, semiconductor) se demuestra muy complejo y únicamente factible de manera aproximada. Por consiguiente, en lo que sigue abordaré el espectro de energías sólo de manera cualitativa.

Mucho más importante resulta cuestionarse cuántos electrones pueden ocupar un nivel energético dado, pues la resistencia de Hall no depende de la energía, sino del número de electrones. Se demostrará que el número de electrones necesarios para ocupar enteramente un nivel energético crece de manera proporcional al campo magnético. La resistencia de Hall deviene independiente del campo magnético y toma un valor cuantizado si los niveles de energía están llenos, de suerte que la corriente puede entonces fluir sin pérdidas disipativas.

## Espectro energético de los electrones de un cuerpo sólido

La energía de los electrones que se mueven libremente se determina a partir de su energía cinética en las tres direcciones espaciales; en principio puede adquirir cualquier valor. Sin embargo, en la resistencia de Hall cuantizada los huecos energéticos desempeñan un papel fundamental: los electrones no pueden tomar determinados valores de energía.

Ese fenómeno admite explicación en el marco de la mecánica cuántica. En esa descripción, a cada electrón se le asocia una onda corpuscular, cuya amplitud está relacionada con la probabilidad de encontrar un electrón en esa posición; la energía del electrón depende de la longitud de onda. Si, por ejemplo, se divide ésta por dos, tres o




**8. EN PRESENCIA DE UN CAMPO MAGNETICO fuerte (dispuesto en dirección perpendicular a la del campo eléctrico), sólo puede haber conductividad en la dirección del campo eléctrico en el caso de que las trayectorias circulares de los electrones se perturben por medio de choques (*estrellas*). Los centros de las trayectorias circulares se desplazan en esas circunstancias con una velocidad de arrastre constante, $v$, en perpendicular a los campos eléctrico y magnético. Gracias a los choques se registra también, sin embargo, un movimiento en la dirección del campo, merced al cual se origina una conductividad finita ($I_0 \neq 0$).**

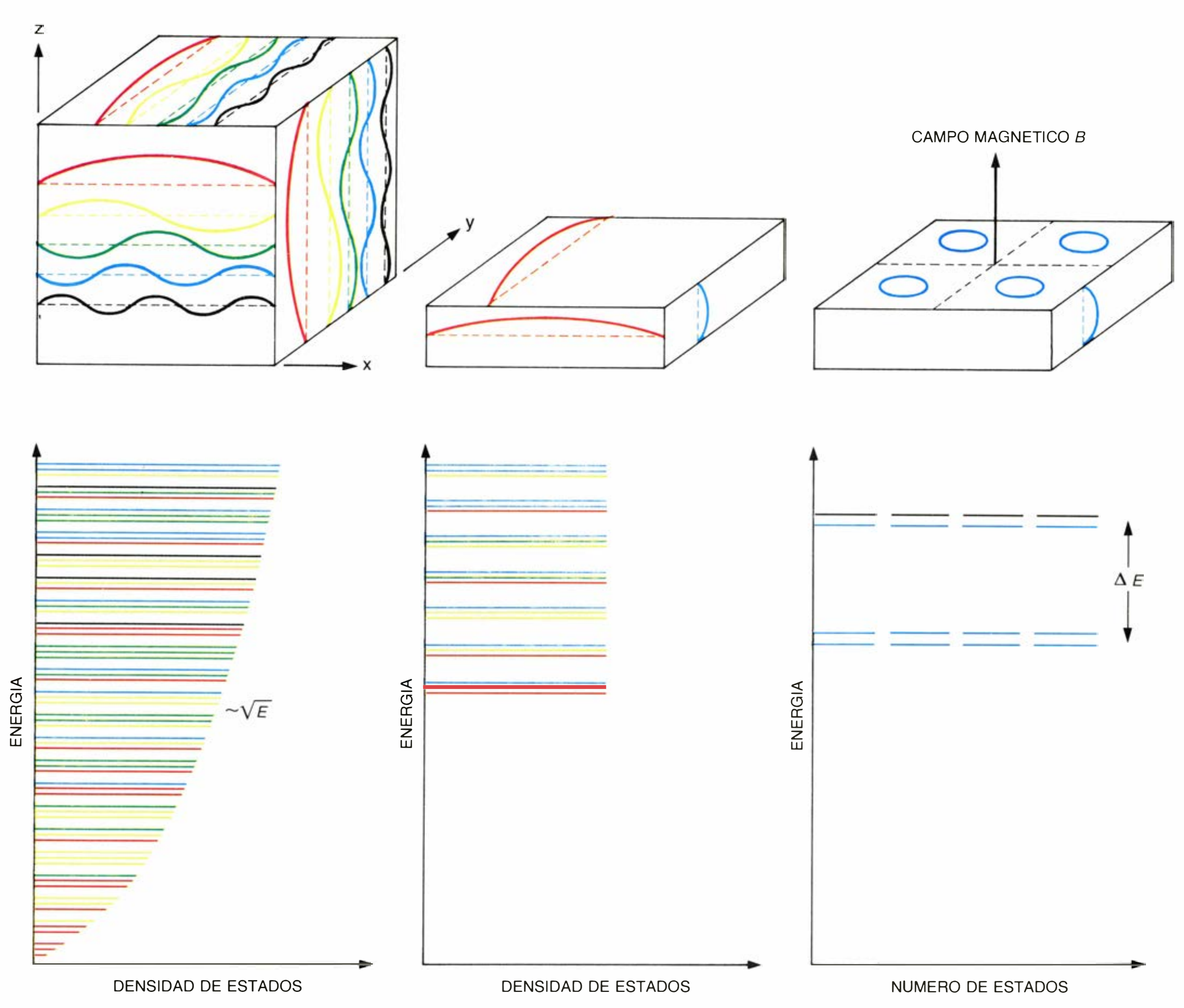
cuatro, la correspondiente energía del electrón queda multiplicada por cuatro, nueve o dieciséis. Tal crecimiento cuadrático de la energía en función de la inversa de la longitud de onda es típico de las ondas corpusculares.

Si encerramos los electrones en una caja, se anula la probabilidad de encontrarlos fuera de la misma. En el contexto de la mecánica cuántica se corresponde ello con exigir que, en el borde de la caja, las amplitudes de las ondas electrónicas sean nulas.

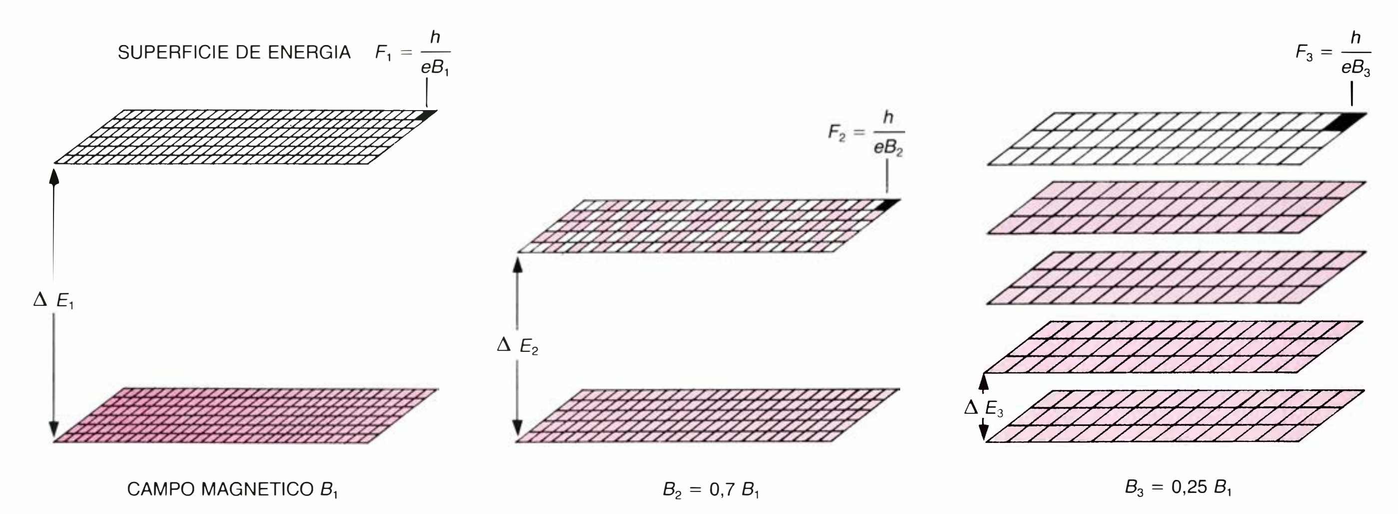
Unicamente ciertas longitudes de onda satisfacen esas condiciones de contorno. El primer electrón que se coloca en la caja toma el más bajo de los valores energéticos permitidos: le corresponde la mayor longitud de onda en cada una de las tres direcciones espaciales. El siguiente electrón debe, en virtud del principio de Pauli, ocupar un estado de superior energía. A medida que se aumente el número de electrones se ocuparán más niveles energéticos, que corresponderán a las diversas combinaciones de las longitudes de onda electrónicas.

En un metal, en el que las medidas de una porción del mismo (por ejemplo, un centímetro de arista) corresponden a las dimensiones de la caja que encierra los electrones, las diferencias de energía entre los niveles aislados resultan inconmensurablemente pequeñas. Se habla de un continuo energético, pues incluso a las temperaturas más bajas que puedan alcanzarse, de una millonésima de grado Kelvin, la energía calorífica de los electrones es miles de millones de veces superior a la diferencia entre niveles de energía vecinos.

Por consiguiente, en una pieza macroscópica de metal resulta imposible




**9. ENERGIA DE LOS ELECTRONES que se mueven libremente en una caja; viene determinada por la longitud de onda electrónica. En el caso tridimensional (*izquierda*) se representan, en las tres direcciones del espacio, las cinco mayores longitudes de onda electrónicas. Al disminuir la longitud de onda crece la energía electrónica; en la figura inferior se representan las energías correspondientes. (Los colores se corresponden con los de las respectivas longitudes de onda. Con tres trazos de color se caracterizan las correspondientes longitudes de onda, una para cada dirección; se trata de las contribuciones a la energía de un electrón.) En la práctica, los niveles energéticos están tan poblados que se habla de un casi-continuo. Los estados energéticos no se distribuyen con densidad uniforme en todo el rango de energías, sino que su densidad aumenta con la energía; lo cual se indica aquí variando la longitud de los trazos. Si la caja en la que están confinados los electrones se aplana (*centro*), se obtiene un gas de electrones casi bidimensional, en el cual la energía de los electrones correspondiente al movimiento perpendicular a la lámina permanece constante. En ese caso la densidad de estados es constante; la distancia energética entre los niveles es tan pequeña que no puede medirse. De aplicarse un campo magnético perpendicular a la lámina (*derecha*) ejercería una fuerte influencia sobre el movimiento libre de los electrones, de tal forma que sólo se permitirían ciertos valores de la energía, separados por una distancia energética medible $\Delta E$. Se obtiene así un espectro energético discreto que recuerda al de la física atómica. El número de electrones de un nivel energético de ese tipo es idéntico al de trayectorias circulares. En la práctica, esa cifra es muy grande, por ejemplo, de 2.417.967.100 para una superficie de un milímetro cuadrado y un campo magnético de 10 tesla.**




**10. ESQUEMA INTUITIVO de la distribución de 180 electrones en los niveles energéticos discretos de un gas de electrones bidimensional, sometido a un campo magnético; en el ejemplo se ilustran los casos de tres valores distintos de la intensidad del campo. Las superficies de energía representadas caracterizan las dimensiones de la muestra; las celdillas, que corresponden cada una a un electrón, tienen un área dada por la cantidad $F = h/eB$. La distancia $\Delta E$ entre los niveles energéticos, y el número de los estados ocupables dentro de un nivel, crecen proporcionalmente con el campo magnético. Tómese el campo magnético $B_1$ de manera que todos los electrones llenen por completo el nivel energético inferior. Si el campo magnético se divide por cuatro ($B_3 = 0,25\ B_1$) se llenarán cuatro niveles. En un campo magnético $B_2 = 0,7\ B_1$ el nivel energético inferior queda ocupado del todo, mientras que el inmediatamente superior queda ocupado aproximadamente hasta la mitad. En un nivel energético parcialmente lleno los electrones pueden trasladarse de una posición ocupada (*roja*) a otra desocupada (*blanca*); de ser ese el caso, en presencia de un flujo de corriente se originan pérdidas energéticas en forma de calor. Cuando la banda de energías está completamente ocupada la corriente puede, por el contrario, fluir sin pérdidas disipativas; la medición de la resistencia de Hall da un valor cuantizado.**

observar los niveles de energía electrónica, en principio discretos. Ahora bien, si la caja de confinamiento no mide más que una millonésima de centímetro, los desdoblamientos de los niveles se tornan mensurables.

Carece de sentido dividir un metal en cubos de una millonésima de centímetro de arista, pues con ellos no pueden efectuarse mediciones de ningún tipo. Sin embargo, sí permite la técnica fabricar láminas finísimas; un material portador aislante impide en ese caso que las láminas se rompan.

De este modo se logra, al menos para *una* dirección de movimiento de los electrones, una cuantización de la energía: valores de la energía claramente separados, correspondientes al movimiento de los electrones en esa dirección. Ese tipo de láminas delgadas suelen denominarse sistemas electrónicos bidimensionales, habida cuenta de que los electrones sólo pueden moverse libremente en un plano.

No obstante, ¿cómo limitar el libre movimiento dentro del plano bidimensional a pequeñas superficies de una millonésima de centímetro de arista, de manera que los niveles energéticos resulten en verdad discretos? Un campo magnético perpendicular a la lámina bidimensional satisface esa condición.

En un campo magnético, los electrones se ven obligados a recorrer trayectorias circulares; están, por tanto, confinados a superficies de valor $F_0$, que se corresponden con la extensión de la trayectoria. Al crecer el campo magnético, $B$, disminuye la superficie que corresponde a una trayectoria circular electrónica; la teoría afirma que, en general, a cada estado electrónico le corresponde una superficie $F_0 = h/(eB)$, donde $h$ y $e$ son las constantes naturales que fijan el valor de la resistencia de Hall cuantizada.

Siendo muy elevado el número de trayectorias electrónicas circulares que se desarrollan en inmediata vecindad (en un campo magnético intenso conviven unos 100.000 millones de electrones en una superficie de un centímetro cuadrado), también lo es el de posiciones electrónicas posibles en un único nivel energético. Si se duplica el campo magnético, en un mismo nivel energético se alojará el doble de electrones.

Puede, por tanto, elegirse el campo magnético de forma tal que, para un número dado de electrones del plano bidimensional, uno de los niveles energéticos quede completamente ocupado. En esas condiciones, los electrones no pueden ya cambiar de lugar por otro vacío, puesto que los huecos más próximos poseen una energía demasiado grande: puede fluir una corriente sin pérdidas disipativas.

Calculando entonces la resistencia de Hall para el correspondiente campo magnético y para la respectiva concentración de electrones, se obtiene la resistencia de Hall cuantizada: cada vez que los electrones llenan un nivel energético [*véase la figura 10*] se mide una resistencia de Hall cuantizada.

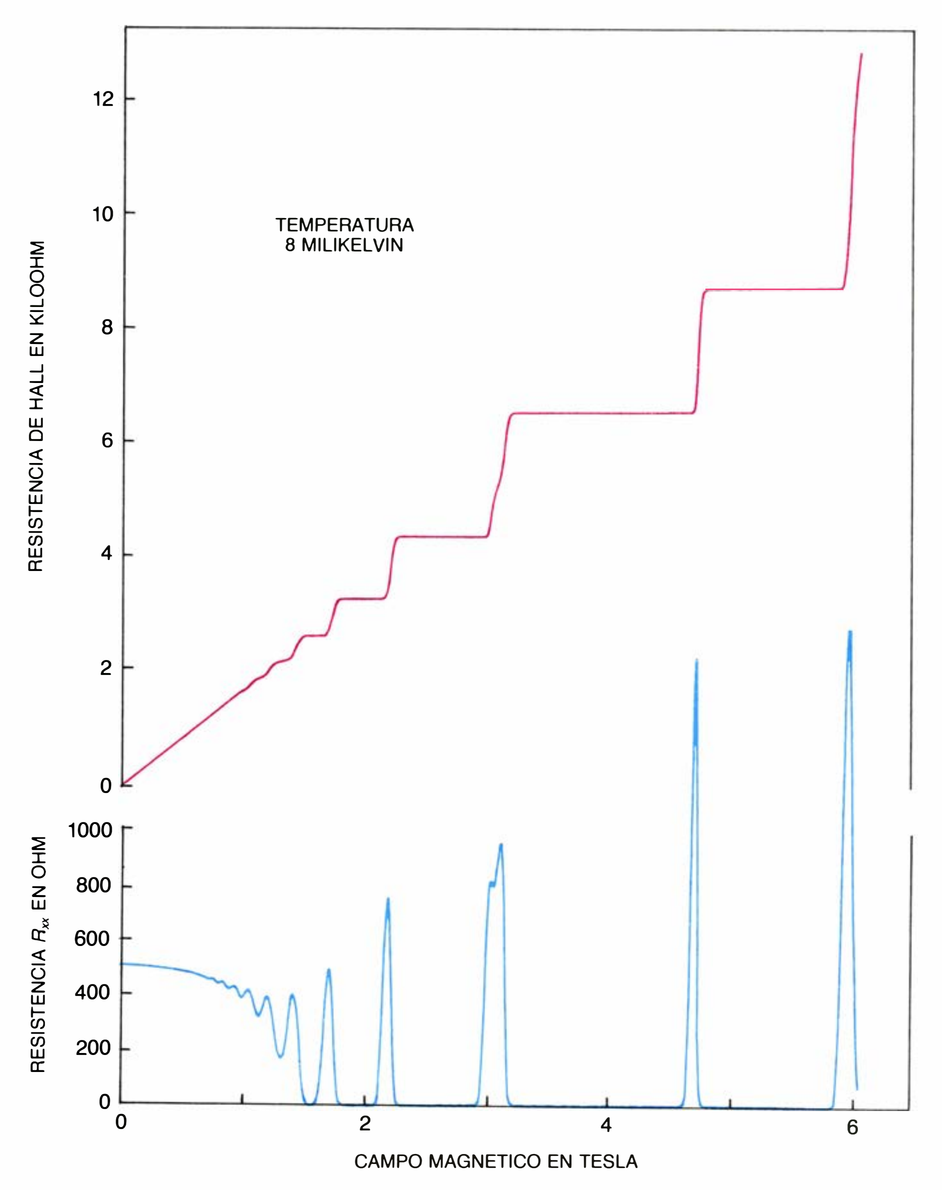
Se observa el efecto cuántico en diversos valores del campo magnético. Por ejemplo si, dado un valor del campo magnético en el que los electrones llenan un nivel, se disminuye ese valor a la mitad, a un tercio o a un cuarto, se ocupan entonces completamente dos, tres o cuatro niveles. Las resistencias de Hall correspondientes serán $h/2e^2$, $h/3e^2$, $h/4e^2$, etcétera.

Al medir las pérdidas disipativas de la corriente (la caída de tensión en la dirección de la corriente) en función del campo magnético [*véase la figura 11*] se observan oscilaciones. La pérdida disipativa se anula en aquellos valores del campo magnético en los cuales los niveles energéticos discretos están enteramente ocupados por electrones. A cada uno le corresponde un valor de la resistencia de Hall cuantizada.

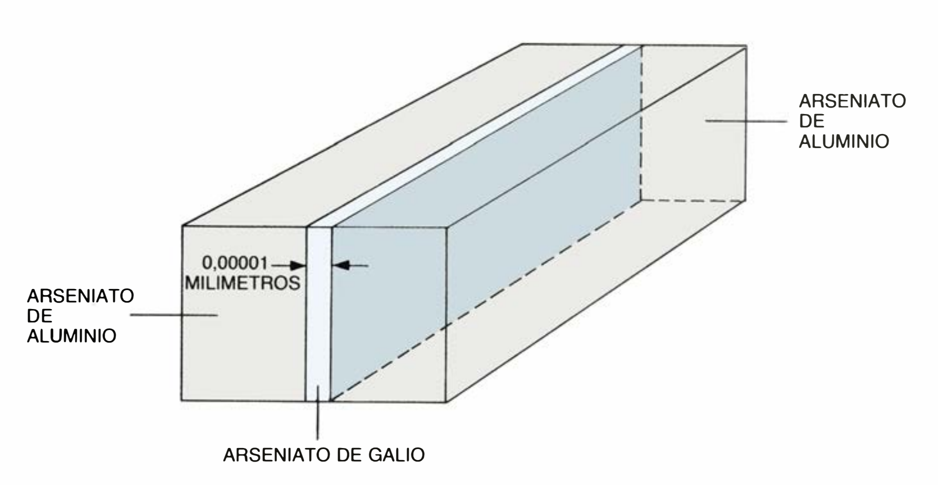
En un experimento con componentes semiconductores reales aparecen siempre perturbaciones de carga eléctrica, que provocan fluctuaciones energéticas en el sistema electrónico bidimensional. En virtud de ello, la corriente de Hall sin disipaciones puede fluir en intervalos enteros de valores del campo magnético, y no sólo en determinados valores: en la gráfica de la resistencia de Hall aparecen "mesetas", que corresponden a los diversos valores de la resistencia cuantizada.

## Sistema electrónico bidimensional

Para observar la resistencia de Hall cuantizada deben satisfacerse tres condiciones. Primera: crear una delgada lámina de electrones de una millonésima de centímetro de espesor, aproximadamente. Segunda condición: un




**11. RESULTADO EXPERIMENTAL de la medición, en el intervalo de campo magnético de hasta 6 tesla, de la resistencia de Hall, $R_H$ (*curva superior*), y de la "pérdida disipativa" de la corriente de electrones, determinada a partir de la caída de tensión en la dirección de la corriente, donde se detecta una resistencia $R_{xx}$ (*curva inferior*). Durante el ensayo, la temperatura se mantuvo a ocho milikelvin sobre el cero absoluto.**




**12. PARA LA OBTENCION de un gas bidimensional de electrones se introduce, con ayuda, por ejemplo, de la llamada epitaxia de rayo molecular, una fina lámina de arseniato de galio de 0,00001 milímetros de espesor en un bloque de arseniato de aluminio, cuya constante reticular es aproximadamente igual a la del arseniato de galio. La afinidad de los electrones por el arseniato de galio es superior a la que muestran por el de aluminio, por lo que prefieren situarse en aquél, formándose una delgadísima lámina de electrones.**

intenso campo magnético, perpendicular a esa lámina ha de obligar a los electrones a describir trayectorias circulares, asimismo de una millonésima de centímetro de diámetro. Como tercera condición, el número de electrones ha de ser precisamente tal que todas las trayectorias circulares posibles, situadas una junto a otra, estén ocupadas por electrones.

La producción de un gas bidimensional de electrones (es decir, la plasmación de la delgadísima lámina de electrones) constituye el punto crucial del experimento sobre el efecto Hall cuántico. Las modernas técnicas de semiconductores ponen a nuestra disposición esas láminas.

Con ayuda de la epitaxia de rayos moleculares cabe fabricar, por ejemplo, un soporte de materiales de distintos tipos en forma de láminas delgadísimas. Si en un material poco afín a los electrones se crea una lámina más inclinada a ellos, los electrones tenderán a reunirse en la lámina, y así se obtendrá el sistema electrónico bidimensional deseado. Una lámina de arseniato de galio rodeada de arseniato de aluminio resulta particularmente apropiada [*véase la figura 12*].

Las primeras láminas bidimensionales de electrones no se obtuvieron, sin embargo, por este procedimiento, sino que (siguiendo otros métodos) se crearon sobre la superficie de silicio semiconductor [*véase la figura 13*]. Por medio de una placa metálica cargada positivamente (la "puerta"), paralela a la superficie del semiconductor, los electrones del silicio se atraen hacia la superficie. Dado que los electrones del material semiconductor no pueden abandonarlo (ello requeriría un trabajo de extracción descomunal) y, por otra parte, los electrones, cargados negativamente, resultan atraídos por la placa metálica, de carga positiva, sobre la superficie del semiconductor se forma una delgada capa de electrones (de suficiente finura para generar niveles energéticos que permitan el movimiento perpendicular de los electrones a esa lámina).

El número de electrones instalados en la lámina superficial es, en el caso ideal, igual al número de cargas positivas de la placa metálica, que puede variarse por medio de una tensión eléctrica entre metal y semiconductor. Cabe así modificar la conductividad de la superficie del semiconductor; cuanto mayor sea la tensión positiva en la placa metálica (tensión de puerta), mayor será la conductividad eléctrica en la

superficie del semiconductor. En esta capacidad de gobernar la corriente eléctrica del semiconductor con ayuda de una tensión se basa el modo de operación del transistor de efecto de campo, la piedra angular de la microelectrónica moderna.

En la práctica, el metal de la fuente no queda separado del material semiconductor por una capa de aire; antes bien, una delgada lámina aislante (dióxido de silicio) actúa de barrera frente a los electrones. La calidad de la superficie fronteriza entre el silicio y el dióxido de silicio es decisiva para el buen funcionamiento de un transistor de efecto de campo; las asperezas u otras perturbaciones de la superficie fronteriza perjudican el transporte de corriente en la lámina superficial.

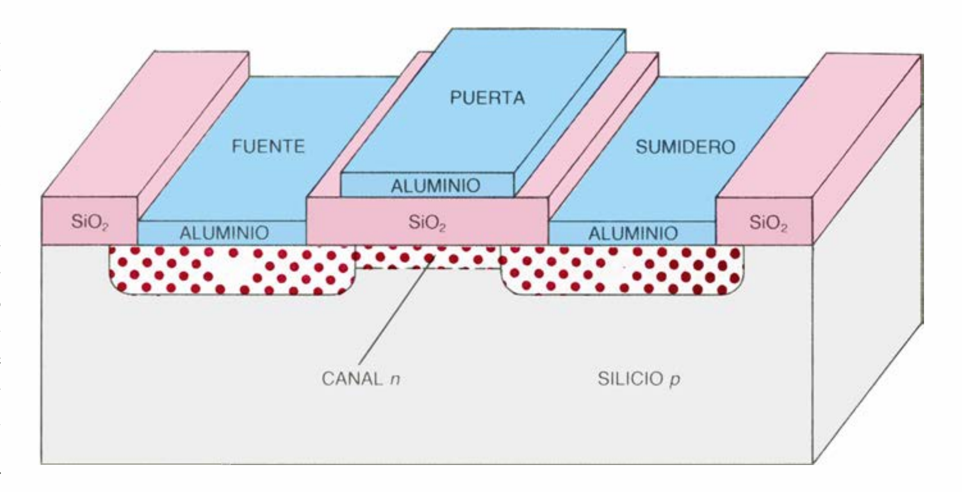
Los procesos técnicos implicados en la fabricación de una buena superficie fronteriza semiconductor-aislante pertenecen al secreto de la producción de componentes semiconductores. Desde hace un par de años, la Unión Soviética ocupa una posición de vanguardia en ese terreno. Las recetas para construir superficies fronterizas entre el dióxido de silicio amorfo, cuyos átomos ocupan posiciones sin orden alguno, y el silicio cristalino, en el que los átomos se disponen en un retículo regular, han resultado hasta la fecha tan eficaces que las grandes empresas de semiconductores se han especializado principalmente en ese sistema.

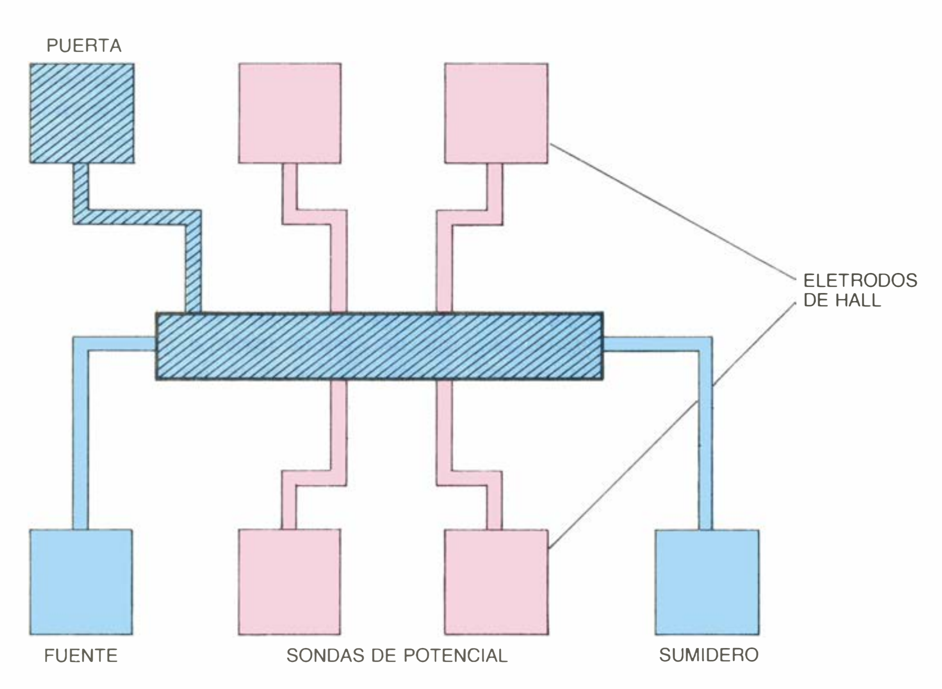
Sin embargo, desde una perspectiva física, las superficies fronterizas de heteroestructuras (que en vez de silicio amorfo emplean un material semiconductor cristalino con propiedades aislantes) son cualitativamente mejores: no presenta aquí el cristal fallo alguno, ni rugosidades la superficie. En los laboratorios de investigación de fabricantes de semiconductores japoneses y norteamericanos se estudian con especial atención las heteroestructuras de arseniato de galio y arseniato de aluminio. Se confía en que las propiedades electrónicas, extraordinariamente buenas, de ese sistema bidimensional de electrones permitan el desarrollo de nuevos componentes y un incremento sustancial de la velocidad de cálculo de los ordenadores.

En ese sistema, el efecto Hall cuántico se observa con singular facilidad. La razón de ello es que la longitud de onda libre sobre la cual puede moverse un electrón sin impedimentos es tan grande que las trayectorias circulares de los electrones se cierran en sí mismas y no se perturban ante la presencia de campos magnéticos relativamente débiles (de grandes radios de trayectoria circular); los niveles energéticos de los electrones en el campo magnético quedan, en consecuencia, cuantizados.

Por regla general, y merced a sus extraordinarias propiedades electrónicas, los sistemas electrónicos en los cuales cobra importancia la resistencia de Hall cuantizada constituyen también interesantes candidatos para la elaboración de nuevos tipos de elementos semiconductores. En los trabajos sobre sistemas bidimensionales de electrones se manifiesta con especial claridad la fluida transición entre la investigación fundamental sobre nuevos efectos físicos y los productos de consumo.

Nadie podía imaginar que las observaciones de fenómenos cuánticos en láminas semiconductoras, efectuadas con el ánimo de analizar la resistencia de Hall cuantizada, habrían de constituir el fundamento sobre el que se asentara el desarrollo de nuevos tipos de semiconductores para láser o de transistores de efecto de campo más rápidos. Probablemente los avances futuros en el terreno de los componentes de semiconductores lleven la impronta de los fenómenos cuánticos.




**13. SECCION DE UN TRANSISTOR de efecto de campo de silicio para mediciones del efecto Hall (*dibujo superior*). La corriente fluye desde la fuente hacia el sumidero. La presencia de una tensión positiva en la puerta genera una delgada lámina de electrones sobre la superficie del semiconductor (canal *n*); de ese modo se gobierna la intensidad de la corriente que discurre entre la fuente y el sumidero. En la figura inferior se muestra un diagrama en planta del transistor de efecto de campo [*véase también la figura 1*]. Junto a los contactos de corriente se han dibujado las sondas de potencial y los electrodos de Hall de medición del potencial. La dirección del campo magnético en las mediciones del efecto Hall resulta perpendicular al diagrama.**

# Juegos de ordenador

## *De cómo un par de programas obtusos pasan por genios en los tests de inteligencia*

A. K. Dewdney

Hay un número muy gastado, pero que provoca indefectiblemente grandes carcajadas, y del que vodeviles y revistas teatrales suelen echar mano con frecuencia: la clásica burla del forzudo de circo. Salen a escena un hombre de tremenda musculatura y su ayudante, una mujer bastante menos dotada en ese aspecto. El hombre, tensos los músculos al máximo, trata de levantar unas enormes pesas. Con esfuerzo sobrehumano, el forzudo logra finalmente alzarlas por encima de la cabeza. Los espectadores aplauden y vitorean; pero los aplausos se tornan risas y chacotas cuando la ayudante coge inocentemente las pesas con una mano y, sin mayores trámites, las retira del escenario.

Dos programas de ordenador dejan en nosotros la misma cómica sensación de anticlímax acerca de la pretendida "potencia" mental que, al parecer, refleja la obtención de altas puntuaciones en el clásico test del cociente de inteligencia (CI). En dos tareas muy frecuentemente utilizadas en tales tests, a saber, la completación de sucesiones numéricas y la percepción de analogías visuales, ambos logran resultados de "genio", o casi. No obstante, poco cuesta comprenderlos, dejándonos estupefactos al percatarnos de lo muy romos que son sus ingenios.

No tengo deseo alguno de ofender a los lectores que se crean generosamente dotados de cacumen, pero sí el serio propósito de poner en solfa el valor de los tests del CI. El propósito declarado del ensayo es la medición de la inteligencia; pocas son las cualidades humanas cuya posesión evoque tanta ufanía, ni tanta ansiedad su carencia. Sin embargo, la noción de inteligencia que las pruebas tradicionales del cociente presuponen está gravemente desenfocada. El razonamiento que sustenta semejante juicio lo expone convincentemente Stephen Jay Gould, de la Universidad de Harvard, en su libro *The Mismeasure of Man*. Hélo aquí, en resumen: los tests tradicionales para medición del CI se fundan en la presunción, errónea y no explicitada, de que la inteligencia, como la fuerza, es una cualidad específica de la fisiología humana, y que puede medirse mediante una serie de tareas graduadas.

Constituye un buen ejemplo de ello la completación de secuencias numéricas. ¿Cuál es el número siguiente en la sucesión 2, 4, 6, 8, . . .? ¿Y en la sucesión 2, 4, 7, 14, . . .? ¿Y en la 1, 2, 6, 24, . . .? El porcentaje de respuestas correctas que demos a una tanda de preguntas similares mediría nuestra "inteligencia general", exactamente igual que un dinamómetro mide el peso que uno puede levantar y, por consiguiente, la fuerza de los músculos de nuestros brazos. Fijémonos en que si los resultados del ensayo deben interpretarse como una medición de la "inteligencia general", tendrá que existir cierta facultad básica, o un reducido conjunto de tales facultades, que proporcione un índice de lo que se entiende por inteligencia general. Dado que la idea misma de inteligencia general presupone una fuerte correlación entre las facultades básicas, el tipo exacto de tarea graduada que se adopte en el test de CI carece, hasta cierto punto, de importancia. Tanto puede servirnos una como otra.

Uno de los programas de CI aquí presentados es adaptación de otro programa más elaborado, debido a Marcel Feenstra, estudiante de Rotterdam. El programa de Feenstra se llama HI Q y resuelve dos tipos de problema que aparecen con frecuencia en los tests ordinarios del CI: la completación de sucesiones y las analogías numéricas. Recientemente, Feenstra puso a prueba el funcionamiento de HI Q, haciéndolo actuar sobre cierto número de tests de CI, tomados de un libro de Hans J. Eysenck, de la Universidad de Londres, titulado *Know Your Own I.Q.* Según parece, el CI del programa HI Q ronda en torno a 160. Cierto es que el experimento no fue precisamente lo que se conoce por ensayo científico en condiciones controladas, pero también lo es que deja pocas dudas de que el programa puntuaría muy alto en auténticas condiciones de examen.

El programa que yo tengo en mente se llama SE Q, y reproduce los resultados de HI Q en las tareas de completación de secuencias numéricas. Los lectores que redacten y hagan funcionar el programa SE Q pueden considerar amplificada su propia inteligencia numérica; por poderes, como si dijéramos. La idea principal del programa es de lo más simple. Cuando nos dan una serie de números y nos piden hallar el que va a continuación en la serie, uno no busca el número directamente. Antes bien, lo que hacemos es buscar la regla que ha generado los números ya dados. Es necesario aquí un inciso de carácter matemático: dada una serie finita de números cualesquiera, existe una infinidad de reglas matemáticas capaces de producirla. La búsqueda se reduce, pues, a hallar una regla sencilla que genere la serie dada.

El programa SE Q solamente atiende a dos tipos de posibles reglas de formación: las aditivas y las multiplicativas. Por ejemplo, para hallar el número siguiente de la sucesión 2, 4, 8, 14, . . . se podría comenzar buscando una regla aditiva; la mejor forma de hallar tal regla es construir lo que yo llamo una pirámide de diferencias [*véase la figura 1*]. En la base de la pirámide se sitúa la sucesión dada; se construye la pirámide de abajo arriba, hallando las diferencias entre cada par de números consecutivos del nivel o fila del número precedente. Así, el primer número de la segunda fila de la pirámide se obtiene a partir de los dos primeros números de la primera fila, a saber 2 y 4. Su diferencia es 2; por tanto, el primer número de la segunda fila es 2. Análogamente, los restantes números de la segunda fila son $8 - 4 = 4$ y $14 - 8 = 6$; la segunda fila está formada por los números 2, 4 y 6. Continuando el mismo proceso en un tercer piso se obtiene una secuencia de sólo dos números, ambos iguales a 2.

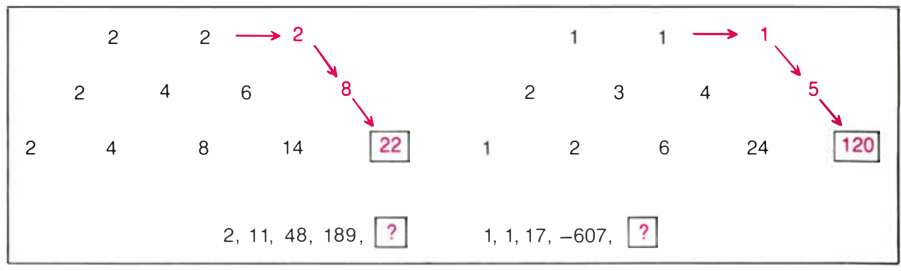
La igualdad de los números de una de las filas de la pirámide es, por así decirlo, la señal de que cese la construcción de la pirámide hacia arriba y comience su extensión lateral. Por ejemplo, supongamos que el tercer número de la tercera fila también sea un 2. Resulta entonces razonable suponer que el siguiente número de la segunda fila se obtenga a partir del número precedente, es decir, 6, sumándole 2: la suma es 8. El número recién deducido en la segunda sucesión puede entonces sumarse al último número dado en la primera sucesión: 14 más 8 son 22, y en efecto, quienes confeccionaron el test

le dan al 22 la calificación de "perfecto". En cuanto en la fila cimera se alcanza una sucesión constante, los nuevos números de la segunda sucesión van infiltrándose pirámide abajo.

Muchísimas de las cuestiones propuestas sobre completación de sucesiones numéricas en las pruebas del CI ceden ante este sencillo procedimiento. Los lectores que tengan con el álgebra algo más que una familiaridad superficial reconocerán en el ejercicio la huella de un polinomio. La serie de valores numéricos de un polinomio, correspondiente a una sucesión de enteros consecutivos, engendra una pirámide de diferencias. Dándole al polinomio suficientes valores, acabará apareciendo en la cima de la pirámide de diferencias una hilera de números idénticos. Restando 1 del número de hileras precisas para formar una pirámide cuya hilera superior sea constante, se tendrá el grado del polinomio. La secuencia 2, 4, 8, 14, . . ., que produce una hilera constante de doses en el tercer nivel de la pirámide, está generada por los valores sucesivos de un polinomio cuadrático (es decir, de segundo grado), a saber, $x^2 - x + 2$.

Desdichadamente, no cabe resolver todas las cuestiones relativas a sucesiones por el método de la pirámide de diferencias. Por ejemplo, la sucesión 1, 2, 6, 24, . . . genera una pirámide de diferencias que tiene en su fila cimera los números 3 y 14. Sin embargo, el rápido crecimiento de sus términos induce fuertemente a pensar en una progresión geométrica. Ahora bien, los términos de una progresión geométrica no están relacionados mediante sumas, sino mediante productos. Así pues, parece razonable formar, a partir de la serie dada, un conjunto de cocientes en lugar de un sistema de diferencias [*véase la figura 4*]. Al tomar los cocientes de los pares sucesivos de términos de la serie 1, 2, 6, 24, . . . se obtiene el segundo piso de una pirámide, formado por la secuencia 2, 3, 4, . . . Esta segunda serie nos señala un abrupto cambio de regla: la tercera fila de la pirámide debe calcularse tomando diferencias, y no cocientes. En efecto, ¿quién dudaría de que la solución deseada exige poner un 5 al final de la segunda hilera? Así, la respuesta pedida es 120, o sea, el producto de 24, el último número de la hilera, por 5.

El programa SE Q, concebido para calcular inductivamente los términos sucesivos de una serie, se esfuerza en construir pirámides, considerando para ello, en cada hilera dada, tanto las diferencias de los pares de números consecutivos como sus cocientes. Más todavía, el programa examina los sucesivos pares de números de cada fila, buscando reglas aditivas y multiplicativas más generales. En una regla aditiva, el primer número de cada par puede estar multiplicado por una constante $k$ antes de que se efectúe la adición habitual; en una regla multiplicativa, la constante $k$ puede sumarse justamente después de la multiplicación habitual. He aquí una pieza sencilla para programadores noveles [*véase la figura 2*].



*1. Se completan las sucesiones numéricas por el método de la pirámide. ¿Podrán los lectores resolver las dos pendientes?*

Unas cuantas fórmulas sencillas, que proporcionan las reglas generales, componen el núcleo de SE Q. Supongamos que la secuencia dada esté asignada ya a las cuatro variables $a(1)$, $a(2)$, $a(3)$, $a(4)$. Para obtener la segunda hilera $b(1)$, $b(2)$, $b(3)$, SE Q prueba a sustituir, o bien una diferencia generalizada, de la forma $b(1) \leftarrow a(2) - k \times a(1)$, o un cociente generalizado: $b(1) \leftarrow [a(2) - k]/a(1)$. En ambos casos, $k$ denota un entero cualquiera de un intervalo predeterminado. El programa ensaya también sustituciones análogas para $b(2)$ y $b(3)$, todas ellas con el mismo valor de $k$: para $b(2)$ tantea con $a(3) - k \times a(2)$, o con $[a(3) - k] / a(2)$, y para hallar $b(3)$ prueba con $a(4) - k \times a(3)$, o con $[a(4) - k]/a(3)$.

La tercera hilera, $c(1)$, $c(2)$, se prepara todavía más sencillamente: SE Q prueba a sustituir en ella simples diferencias: $c(1) \leftarrow b(2) - b(1)$ y $c(2) \leftarrow b(3) - b(2)$, o simples cocientes, $c(1) \leftarrow b(2)/b(1)$ y $c(2) \leftarrow b(3)/b(2)$. Al parecer, es raro que en los tests del CI se propongan cuestiones de prolongación de sucesiones más complejas de las que estas fórmulas consienten.

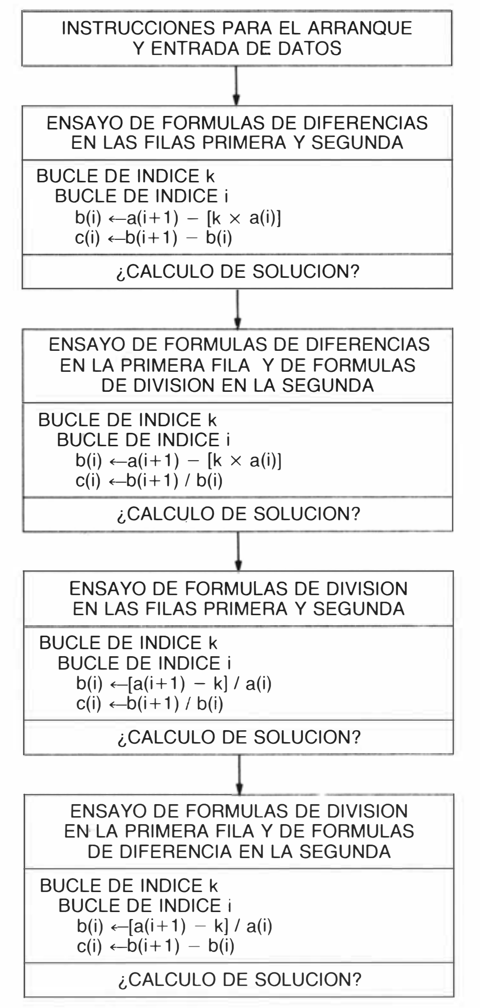
Para ir generando una pirámide, SE Q ensaya cada una de las sustituciones generalizadas correspondientes al sistema de variables $b$ con cada una de las sustituciones sencillas correspondientes al conjunto de variables $c$. Por consiguiente, desde el punto de vista conceptual, SE Q se compone de cuatro segmentos principales. Cada segmento consiste en un bucle, en cuyo interior se halla una combinación de fórmulas de sustitución. Por ejemplo, un tal segmento del programa comienza por aplicar las tres fórmulas de división generalizada, cuya expresión es del tipo $b(1) \leftarrow [a(2) - k]/a(1)$, a fin de calcular los valores $b(1)$, $b(2)$ y $b(3)$ que componen la segunda hilera de la pirámide. El segmento de programa aplica entonces las dos sencillas fórmulas de diferencia, que son del tipo $c(1) \leftarrow b(2) - b(1)$, al objeto de calcular $c(1)$ y $c(2)$, que componen la tercera hilera. El conjunto completo de fórmulas utilizadas en este segmento está integrado en lo que podríamos llamar un "bucle pruébalotodo", en el cual se tantea con distintos valores de $k$. Feenstra recomienda que se permita a $k$ tomar todos los valores enteros desde $-5$ hasta 5.

Cada vez que en el interior del bucle se calculan nuevos valores de $c(1)$ y $c(2)$, se comprueba si son iguales o no. Si resultan ser iguales, su valor común se almacena en una variable llamada $c$, y el correspondiente valor de $k$ se guarda en una variable llamada $kk$. Justamente a continuación del bucle viene un grupo de instrucciones que construyen, a partir de $c$ y $k$, la solución que continúa la sucesión de partida (si es que se ha hallado alguna). En el ejemplo que estoy describiendo, el nuevo miembro de la segunda hilera, $b(4)$, se obtiene sumándole $c$ a $b(3)$. La solución, $a(5)$, se obtiene entonces multiplicando $a(4)$ por $b(4)$ y sumándole $kk$ al producto.

Bastan, pues, dos instrucciones a la salida del bucle para generar una solución a partir de una búsqueda con éxito en el interior del bucle. Las instrucciones adecuadas para cada bucle dependen de las fórmulas utilizadas en él, y por mi parte dejaré al cuidado de quienes redacten SE Q la tarea de descubrir por sí mismos tales instrucciones. Utilícese una pizca de álgebra para aislar la variable que interese. En cada caso, si uno de los bucles del programa descubre una solución, tiene que haber una instrucción que ordene imprimirla. El programa puede entonces saltarse todos los demás bucles y detenerse, o bien ejecutarlos todos, en un esfuerzo

por descubrir más de una solución. Feenstra, ejecutándolos en su totalidad, ha detectado varias cuestiones "mal planteadas" en los tests de CI, es decir, cuestiones que admiten más de una solución. En el caso de que ninguno de los cuatro bucles logre dar con una solución, es razonable añadir, al final del lote completo, un enunciado de salida adicional. El mensaje que tal enunciado ordene emitir puede adaptarse al gusto de cada cual; quienes deseen infundir en sus programas un poco de personalidad pueden mandarle imprimir "¡Socorro!".

Podemos poner a prueba el SEQ con cuestiones tomadas de los tests de muestra que figuran en varios libros fácilmente disponibles. En el libro de Eysenck hay ocho exámenes completos, que se afirma permitirán al lector determinar su cociente intelectual. Dichos tests contienen distintos tipos de tareas que figuran en las pruebas de CI ordinarias, entre ellas, cuestiones relativas a números, letras, o palabras ausentes, a determinar elementos u objetos que se encuentran fuera de lugar, palabras revoltijadas y analogías de




*2. Diagrama de flujo correspondiente a SEQ*

tipo visual [*véase la figura 3*]. Por lo común, dentro de cada una de las categorías principales puede haber varios subtipos. Por ejemplo, en el libro de Eysenck hay tres clases de cuestiones relativas a un número ausente, a saber, los problemas ordinarios de continuación de sucesiones numéricas, ya descritos, y otras dos clases más, de las que son típicos los siguientes ejemplos:

| | | |
|---|---|---|
| 164 | (225) | 286 |
| 224 | (—) | 476 |

| | | |
|---|---|---|
| 8 | 3 | 21 |
| 6 | 5 | 25 |
| 12 | 2 | — |

En cada caso, el aspirante a genio tiene que deducir el número que falta, de acuerdo con ciertas reglas de formación que crea percibir. El programa HIQ se ocupa de las cuestiones de este tipo mediante procedimientos inspirados en el mismo tipo de fórmulas que las utilizadas por el programa de completación de sucesiones. Por mi parte, animo a los lectores para que prueben su mano en ellas; el mes que viene daré las soluciones de los dos problemas anteriores, así como los planteados en las figuras 3 y 4.

Aunque HIQ tan sólo resuelve uno de los principales tipos de cuestiones propuestas en los tests de CI, también podemos mecanizar cuestiones de otros tipos. En realidad, hace más de 20 años que Thomas G. Evans preparó un programa que resuelve analogías de carácter visual; programa que formaba parte de una tesis doctoral defendida en el Instituto de Tecnología de Massachusetts. Aunque ello pueda hacer pensar que se trata de un programa de gran dificultad, la verdad es que las ideas esenciales del programa de Evans se entienden fácilmente.

Las analogías visuales que el programa de Evans resuelve se ajustan todas a una misma pauta, a saber, la figura *A* es a la figura *B* como la figura *C* es, por ejemplo, a una de cuatro figuras, presentadas como posibles respuestas. Para seleccionar la figura análoga, el programa empieza por determinar primero un conjunto sencillo de reglas que puedan transformar la figura *A* en la figura *B* [*véase la ilustración 4*]. Seguidamente repite el procedimiento con la figura y con cada una de las cuatro hipotéticas respuestas; en cada caso genera un conjunto de reglas que puedan transformar la figura *C* en la hipotética solución. La figura obtenida por las reglas de transformación que más se asemejen a las que llevan de la figura *A* a la *B* es la seleccionada como solución.

En esencia, lo que hace el programa de Evans es repetir una misma operación cinco veces. En cada ocasión actúan como datos de entrada dos figuras, a saber, una figura fuente y una figura destinataria. El programa comienza por enumerar las relaciones espaciales entre las partes de la figura fuente; seguidamente enumera las relaciones espaciales entre las partes de la figura destinataria. En ambas descripciones se consideran únicamente tres relaciones espaciales, *encima de, a la izquierda de* y *dentro de*. Finalmente, el programa describe cómo cambiar las partes de la figura fuente en las partes de la figura destinataria mediante uno cualquiera de cuatro procedimientos fundamentales: cada parte puede cambiar de forma, girarse, copiarse por simetría o eliminarse.

Supongamos que las figuras *A*, *B* y *C* constan cada una de tres partes: un círculo, un cuadrado y un triángulo. En las figuras *A* y *B* el programa puede rotular con una *a* los triángulos, con una *b* los cuadrados y con una *c* los círculos; en cambio, no hace intento alguno de rotular de igual modo las partes que componen la figura *C*. Asigna, en su lugar, las letras *x*, *y* y *z* a las tres partes de la figura *C*. Seguidamente prepara para la pareja de figuras *A* y *B* una descripción tabular de las tres partes de que constan y otras cuatro descripciones más, correspondientes cada una a un emparejamiento de la figura *C* con cada hipotética solución. En estas cuatro últimas tablas se utilizan exclusivamente los rótulos *x*, *y* y *z*.

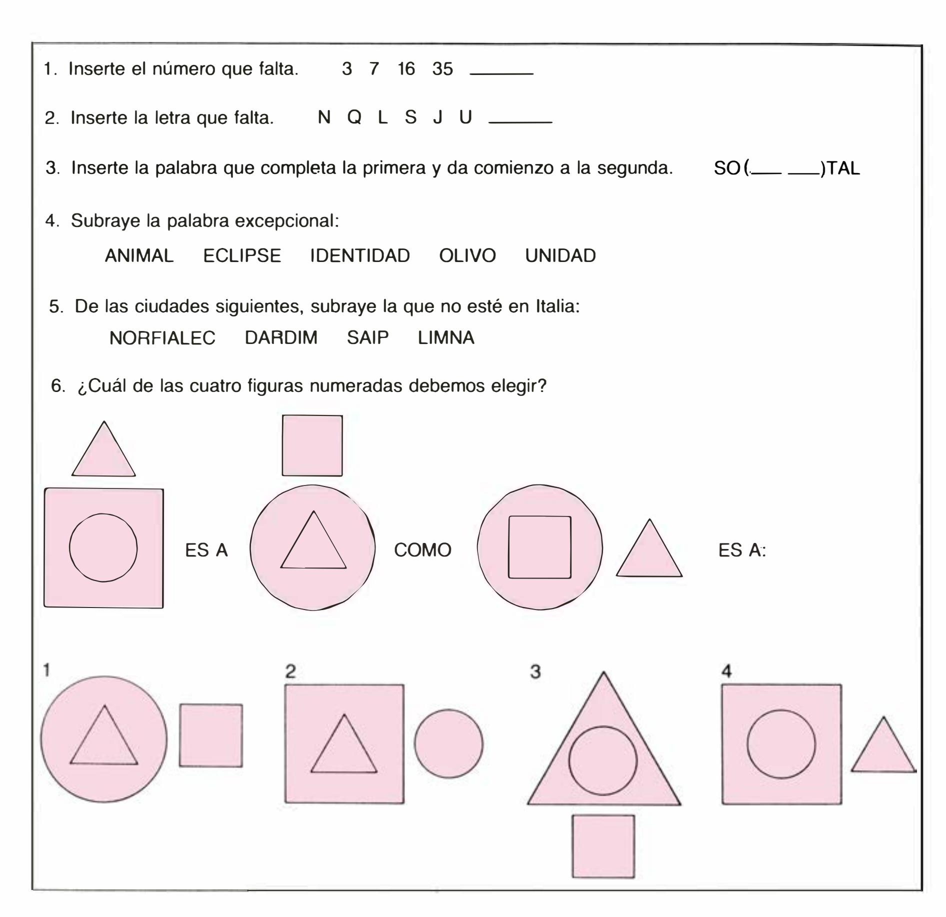
La última operación del programa de Evans consiste en sustituir *x*, *y* y *z* por *a*, *b* y *c* de todos los modos posibles. Dado que *x*, *y* y *z* sólo pueden permutarse de seis modos, es preciso ensayar seis sustituciones. Puede que una de las sustituciones convierta la descripción tabular del par de figuras *A* y *B* en la correspondiente descripción tabular de la figura *C* emparejada con una de las potenciales soluciones. La figura que forme pareja con *C* es la solución. Sin embargo, aunque no llegue a descubrirse una correspondencia exacta, el programa puede valorar el éxito relativo de una analogía y elegir incluso en tal caso la sustitución que dé resultado óptimo.

Patrick Henry Winston describe el programa de analogías visuales en su libro *Artificial Intelligence*. En él, Winston afirma que el programa "funciona bien", y atribuye su éxito al empleo de un marco referencial eficaz donde representar los conocimientos concernientes a las figuras geométricas consideradas en el programa. Por ejemplo,

en lugar de especificar cómo cambian de una figura a otra las relaciones de posición tales como *encima, a la izquierda* y *dentro,* el programa podría haber intentado describir por qué procedimiento queda transformada una figura del primer dibujo en otra del segundo. Tal programa sería extraordinariamente fastidioso, por no decir ineficaz, porque tendría que revisar un número mucho mayor de potenciales sustituciones que el de Evans. En efecto, uno de los temas principales de la inteligencia artificial es la búsqueda de representaciones idóneas; con mucha frecuencia, ésta es la clave que consiente a un ordenador remedar algún aspecto de la capacidad de los humanos para resolver problemas.

Por analogía, la representación de objetos en la mente humana resulta también materia de buena parte de los análisis y discusiones de los especialistas en procesos cognitivos. En tal contexto es frecuente que el estudio de la inteligencia artificial se justifique como un intento de exhibir una "demostración de existencia" para las descripciones mecanicistas de las facultades humanas. Así, prosigue tal argumentación, si se logra que un programa simule algún aspecto de la conducta humana, la representación de la conducta adoptada por el programa podría cuando menos hacer el papel de la representación subyacente adoptada por el cerebro. Empero, en muchas ocasiones, las más felices simulaciones de tal conducta no permiten atisbar, o apenas, de qué modo realizan las personas esas mismas cosas. Que uno sepa, puede que no haya relación alguna entre el funcionamiento de los programas de CI ideados por Feenstra o Evans y los procedimientos por los cuales las personas resuelven esos mismos problemas. Es de presumir que la inteligencia humana despliegue estrategias más generales al atacar problemas concretos.

Con lo cual queda del todo cerrado el círculo que me hace retornar a la consideración de la inteligencia humana: ¿qué es y de qué modo se la mide? Como ya hice notar, Stephen Jay Gould ha caracterizado al CI como una "malmedición del hombre". Sus críticas aportan amplia documentación probatoria de dos de las principales falacias subyacentes a la noción misma de medición de la inteligencia: la reificación acrítica de una noción abstracta y la gradación de la noción cosificada sobre una escala unidimensional. En gran medida, el propio lenguaje da cuenta de nuestra tendencia a convertir en cosas lo que en el mejor de los casos no serían sino nebulosas abstracciones. Una vez más, en cuanto nos hemos persuadido a nosotros mismos de que estamos manejando una cosa, nuestro reflejo es medirla.



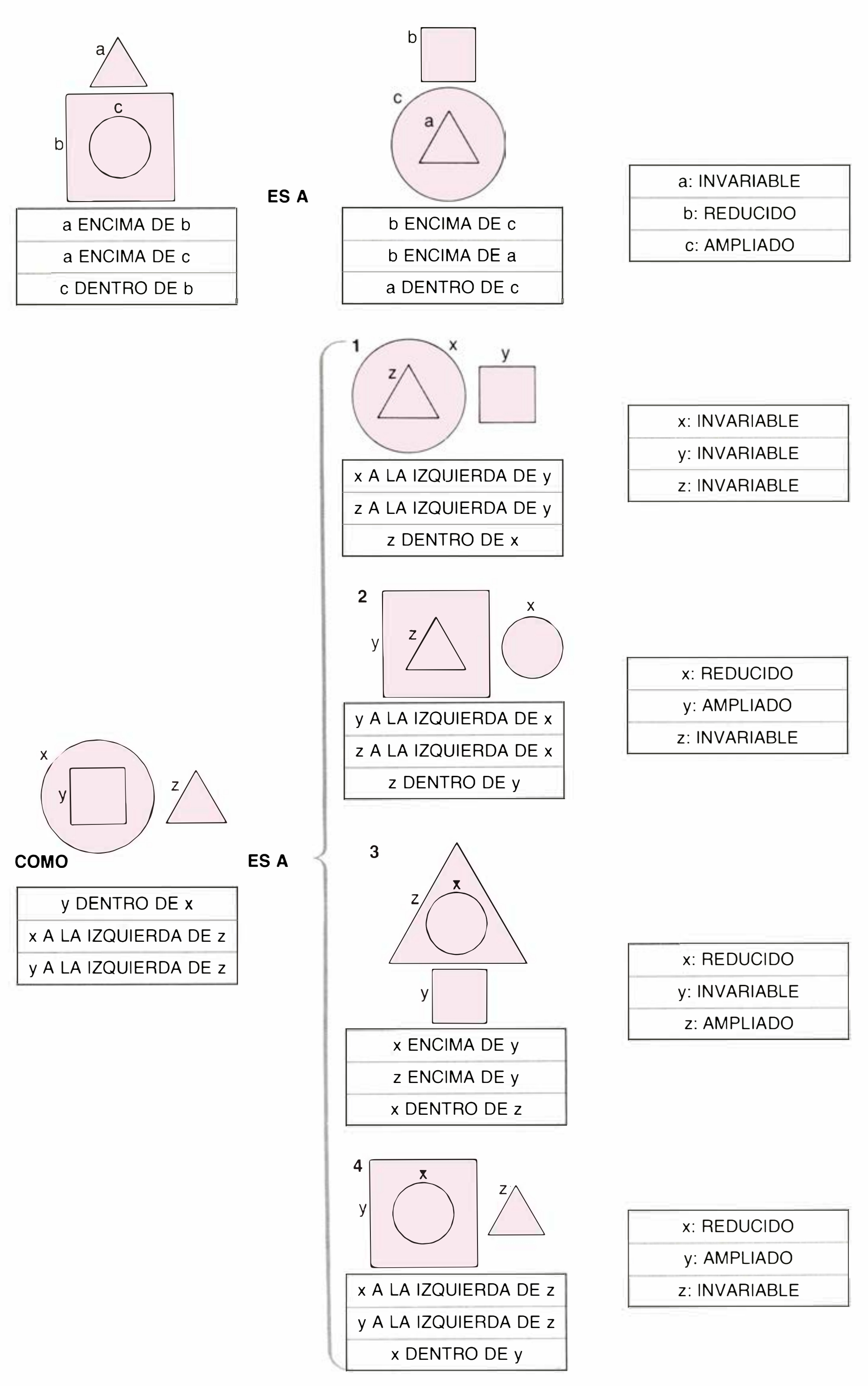

*3. Minitest de inteligencia general basado en cuestiones del libro* Know Your Own I.Q., *de Hans J. Eysenck*

Sucumbimos a la segunda falacia al demandar además una única medida numérica, una gradación lineal. Ansiamos reducir fenómenos complejos a una única escala. Tal proceder ha dado excelentes resultados en física, pero también los ha dado muy mediocres en ciencias sociales. Los tests del cociente de inteligencia constituyen ejemplos al caso; son al siglo XX lo que la craneometría al XIX. En ambas ocasiones, grupos raciales completos se han visto sistemáticamente mal medidos, y no sólo porque la propia medición fuera casi del todo absurda, sino también a causa de los sesgos (deliberados o inconscientes) del propio proceso de medida.

Gould ataca vigorosamente el determinismo biológico, que sostiene que la conducta humana está determinada por genes, y previene contra la consideración de que las facultades de nuestros cerebros sean producto directo de la selección natural. "Nuestros cerebros son ordenadores de enorme complejidad", escribe. "Supongamos que instale yo un ordenador, mucho más simple, para llevar la contabilidad de una fábrica. Tal ordenador podría realizar también otras muchas tareas, mucho más complejas que aquélla para la que se puso, y desprovistas de toda relación con ella. Estas facultades adicionales serían consecuencia ineluctable del diseño estructural, y no de adaptaciones directas. Aunque hubo también razones para construir nuestros ordenadores orgánicos, inmensamente más complejos, éstos poseen un repertorio casi aterrador de otras facultades, entre las que se cuentan, sospecho, casi todas las que nos hacen humanos."

Con esta última metáfora, Gould ha metido el dedo en lo que a mi parecer tiene de más inquietante un programa de ordenador capaz de puntuar a nivel de genio en una prueba de inteligencia. ¿Constituye la puntuación alcanzada en el test una medida de la inteligencia del ordenador? Y si no lo es, ¿cómo se procede para medir la inteligencia de un ordenador, sea éste de plástico y silicio, o de carbono y tejido orgánico? Respuesta: sometiendo un programa del CI a una batería de tests probablemente no.

Las reglas de Golomb, que fueron tema del artículo de febrero pasado, resultaron ser el más arduo de los proyectos a que se han enfrentado los lectores hasta la fecha. Por así decirlo, muchos fueron los llamados y pocos los




*4. Resolución de las analogías visuales en la tesis doctoral de Thomas G. Evans*

elegidos. Varios lectores pretendieron incluso reclamar un premio de 100 dólares ofrecido por el inventor de las reglas, Solomon W. Golomb, de la Universidad del Sur de California.

Una regla de Golomb de *n* marcas constituye la más corta dotada de las siguientes propiedades: lleva grabadas *n* marcas (incluidas las de los extremos), situadas en puntos correspondientes a valores enteros, y permite medir el máximo número posible de longitudes enteras comprendidas entre 1 y la longitud de la regla. La medición de cada longitud no ha de poder hacerse más que de una manera. La única forma lícita de medir una distancia es que corresponda al intervalo entre algún par de marcas. La regla no es golombina si hay dos o más pares de marcas separadas por una misma distancia.

A la hora de redactarse aquel artículo no se conocían reglas golombinas de más de 13 marcas, y la más corta de las reglas de 15 marcas conocidas a la sazón tenía una longitud de 155 unidades. Poco después, Douglas S. Roberston, de la Administración Oceánica y Atmosférica de los Estados Unidos, descubrió una regla de 15 marcas más corta, cuya longitud era sólo de 153 unidades. Después, James B. Shearer, del Centro de Investigación Thomas J. Watson de la IBM, programó un ordenador desocupado a fin de que efectuase una búsqueda exhaustiva de reglas golombinas; el ordenador ha producido hasta ahora reglas de Golomb de 14 y 15 divisiones. La regla golombina de 14 marcas tiene una longitud de 127 unidades, y sus marcas están situadas en 0, 5, 28, 38, 41, 49, 50, 68, 75, 92, 107, 121, 123 y 127. La regla de 15 marcas tiene una longitud de 151 unidades, y sus marcas están situadas en 0, 6, 7, 15, 28, 40, 51, 75, 89, 92, 94, 121, 131, 147 y 151. Shearer escribe que ahorró mucho tiempo de cómputo exigiendo que la marca central de la regla se situara a la izquierda del centro geométrico.

Fue otro problema propuesto por Golomb el que suscitó las reclamaciones del premio de 100 dólares. Las recibidas hasta la fecha son inválidas, pues se basan en una comprensión errónea del problema. Golomb me ha urgido a que aclare las cosas, volviendo a enunciarlo. Hállense dos reglas distintas (sean de longitudes mínimas o no) que tengan ambas el mismo número de marcas, para algún número mayor que 6, capaces ambas de medir el mismo conjunto de distancias. Como antes, en ninguna de las reglas puede haber ninguna distancia medida por más de un par de marcas. No se cuentan como distintas las reglas que sean copia simétrica una de la otra, como por ejemplo la de marcas en 0, 2, 5, 6 y la 0, 1, 4, 6. Hay una infinidad de parejas de reglas conocidas, casi todas ellas no minimales, que resuelven el problema análogo al de Golomb, para seis marcas. Por ejemplo, una de tales parejas tiene marcas en 0, 1, 4, 10, 12, 17 y en 0, 1, 8, 11, 13, 17. Se trata de reglas esencialmente diferentes, que no son una reflejo de la otra; ambas permiten medir todas las distancias comprendidas entre 1 y 17, a excepción de 14 y 15. El premio será para la primera persona que descubra una tal pareja de reglas con más de seis marcas cada una.

La red de asesoramiento, para ayudar a los noveles en programación a llevar adelante los proyectos que se vayan proponiendo en esta sección, ha tropezado con una dificultad imprevista: asesores los hay por centenares, lo que no hay son bisoños. Quizás el nombre de novato haya sido poco feliz. ¿Habrá servido para desanimar a quienes, sin embargo, no programan con mucho éxito? Es el momento de enviarme una postal con su nombre, dirección y teléfono, remitiéndola a esta revista.

# Taller y laboratorio

*Teoría y práctica de la percepción de colores en un retículo blanco y negro*

Jearl Walker

En 1965, Celeste McCollough, del Oberlin College, informó de un desconcertante fenómeno: una situación en la que el sistema visual humano impone el color en un retículo blanco y negro. Primero se observa un retículo de líneas negras con un color intercalado. Luego se mira a un retículo blanco y negro de igual orientación y espaciado que el primero. Entonces las líneas blancas parecen estar tintadas del color complementario al que había en el primer retículo.

La aparición del color ficticio sorprende por varias razones. Se produce sólo si el segundo retículo tiene la misma orientación y espaciado que el primero. Además, aunque se debilite con el tiempo, se manifiesta aun cuando se retrase la observación del segundo retículo durante horas, días o semanas. (Su intensidad, tras un retraso, depende parcialmente de los hábitos alimentarios y de sueño del observador.)

Una persona cuya visión de los colores sea normal debe ser capaz de percibir la coloración de McCollough observando los retículos que se presentan en las páginas que siguen. Comiéncese con el retículo coloreado de la figura 1. Se observará que entre las líneas negras horizontales están intercaladas líneas verdes y entre las verticales otras de color magenta. La ilustración debe estar bien iluminada. La mirada no hay que fijarla, sino desplazarla de modo que las diferentes zonas coloreadas se vean aproximadamente durante el mismo tiempo. Pasados cinco minutos o más, examínese los retículos sin colores de la figura 3. La demostración dará mejor resultado si en este caso la iluminación es baja. Ahora deberán percibirse unos colores débiles y no saturados, sobrepuestos a las líneas blancas.

Los colores que se perciben están asociados a la orientación de las líneas negras. Mientras que las horizontales estaban antes intercaladas en verde, ahora lo están de magenta. Las verticales cambian de color exactamente al revés. Girando la ilustración 90 grados puede comprobarse que los colores están realmente asociados a la orientación de las líneas: el magenta aparecerá nuevamente con las rayas horizontales y el verde con las verticales. No importa la extensión de las zonas; el color se asocia a la orientación.

No hay que confundir el efecto McCollough con el fenómeno conocido como postimagen negativa. En este último el color se desvanece rápidamente en comparación con la coloración de McCollough. Para evidenciar una postimagen negativa fíjese la vista en un campo de color verde durante unos cinco minutos y a continuación en una superficie blanca e indistinta. Durante un corto intervalo de tiempo se percibirá un magenta. Pero si la contemplación de la superficie blanca se retrasa 10 minutos o más, la postimagen no llega a aparecer.

Cuando miramos a una superficie blanca de un modo normal, todos los sensores cromáticos de la retina envían sus señales al cerebro. En algún lugar del trayecto, esas señales son analizadas por comparación de pares de colores complementarios que pugnan entre ellos. Si de una pareja se ve la misma cantidad de cada color, el resultado es ningún color. Por ejemplo, el verde y el magenta compiten; si al ojo llegan luces verde y magenta de la misma intensidad, el resultado es un empate y percibiremos blanco, o sea, una iluminación sin color.

La postimagen normal suele atribuirse a la fatiga de algunos de los sensores cromáticos de la retina. Por ejemplo, si observamos un color verde, los sensores responsables del envío de esa señal al cerebro se hacen menos sensibles. Supongamos que miramos a una superficie blanca inmediatamente después de haberse fatigado los sensores verdes. Aunque en el ojo penetren verde y magenta con la misma intensidad, la débil respuesta de los sensores verdes permite que prevalezca el magenta y percibimos magenta en la zona blanca. Cuando la fatiga haya pasado, la competencia entre los dos colores volverá al empate y veremos blanco.

La coloración de McCollough difiere asimismo de la postimagen normal en que las drogas pueden alterar su duración, lo que sugiere que depende de los neurotransmisores situados en el trayecto visual que une la retina con el cerebro. C.C.D. Shute, de la Universidad de Cambridge, ha informado de que la cafeína acelera la desaparición del efecto, mientras que el té de ginseng fresco (no la variedad instantánea) la retrasa. Bamidele O. Amure, de la Universidad de Cambridge, ha escrito que la nicotina lo prolonga. Además, D.M. MacKay y Valerie MacKay, de la Universidad de Keele, Inglaterra, han comprobado que la debilitación del efecto puede atrasarla el sueño y que la intensidad de la coloración puede depender notablemente de lo bien que duerma el observador antes de mirar al retículo coloreado.

He realizado algunos experimentos acerca del efecto McCollough preparando varios retículos de rayas negras y fotocopiándolos con una máquina que retuviera el trazo negro continuo de las rayas. Luego coloreé los espacios entre éstas. La mayoría de las veces ensayé con un retículo de rayas negras horizontales intercaladas de verde. Lo miraba durante unos cinco minutos y después esperaba cinco más antes de mirar a otro no coloreado. Este retraso me aseguraba que no aparecía el coloreado debido a la postimagen negativa. De cualquier modo, la postimagen no me creó ningún problema, ya que no fijaba la vista. Más dificultad me supuso la necesidad de esperar a que la coloración de McCollough se disipara completamente entre experiencias. De vez en cuando tardaba hasta un día en desaparecer.

¿Necesitaba bordes rectos el efecto McCollough? Me hice un conjunto de retículos iguales a los anteriores, salvo que sustituí las rayas negras por hileras de pequeños puntos negros. La coloración no se manifestó. Cambié los puntos por hileras de círculos negros rellenos, pero la coloración siguió sin presentarse.

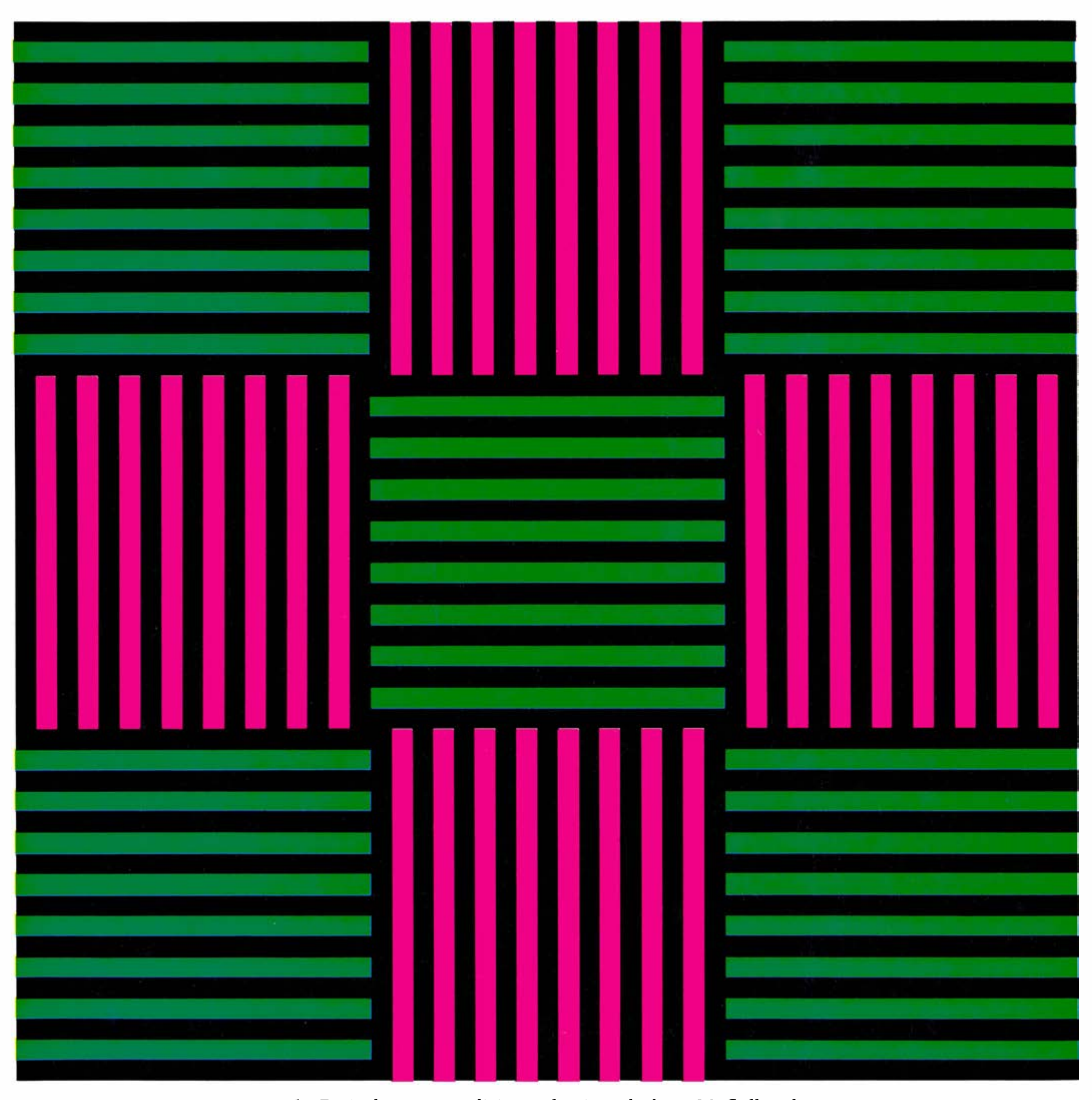
Pensé que acaso el espaciado del retículo alterase la coloración. Acondicionando los ojos al retículo original verde y negro, miré después a otro sin color con la misma orientación, pero cuyas líneas blancas y negras estaban

trazadas a una escala que reducía su tamaño a la mitad, aproximadamente. No apareció coloración.

Me pregunté qué extensión del campo visual debe ocupar el retículo original para que los ojos se acondicionen correctamente al efecto del color. Tapando casi todo el retículo original con papel opaco, miré fijamente a la pequeña superficie aún visible y después examiné el retículo sin color, sin tapar parte alguna del mismo. No se presentó coloración. Evidentemente, para que el efecto McCollough acudiese había que ver una porción suficientemente grande del retículo original para reconocerlo como tal.

Seguidamente invertí el procedimiento, mirando al retículo original completo y examinando después sólo una parte reducida del no coloreado, en su mayoría cubierto de papel opaco. Apareció el magenta pálido. De hecho, parecía seguir allá aunque tapase toda la superficie salvo un pequeño segmento de una o dos rayas negras. Ciertos estudiosos del efecto McCollough han sugerido que si las rayas de un retículo están muy espaciadas en el campo visual, el sistema visual se centra en los bordes de las rayas. Pero si las rayas están estrechamente apretadas en el campo visual, el sistema se centra en la periodicidad de las rayas. Al examinar yo la pequeña porción del retículo sin color, la coloración McCollough estaba probablemente asociada a los bordes más que a la periodicidad.

En 1974 Charles F. Stromeyer III, de la Universidad de Stanford, expuso que la coloración de McCollough se perci-



*1. Retículos que condicionan la vista al efecto McCollough*

| | |
|---|---|
| Verde | Magenta |
| Azul verdoso | Rojo |
| Azul | Naranja |
| Violeta | Amarillo |

*2. Colores complementarios*

bía mejor si el retículo sin color se examinaba con luz tenue. La coloración puede verse aun cuando la iluminación sea tan débil que los conos de la retina, que son los responsables de la información cromática, no actúen. Estos resultados los verifiqué acondicionando mis ojos al retículo negro y verde bajo una luz intensa, disminuyendo luego su intensidad. A medida que mis ojos se ajustaban a la luz tenue, la coloración magenta se acentuaba más, aunque no pudiera percibir colores en ningún otro sitio de la habitación.

Escribía, asimismo, Stromeyer que el espaciado del retículo puede determinar la coloración. Sea un observador que mire primero a un retículo verde y negro de espaciado ancho. Supongamos que mire luego a un retículo magenta y negro de espaciado estrecho. Ambos retículos tienen la misma orientación. Supongamos que pasado un tiempo mire a unos retículos sin color de la misma orientación que los anteriores. Si uno de estos retículos sin color tiene el mismo espaciado que el primero, percibirá magenta; pero si el espaciado es el del segundo, percibirá verde. El observador ha almacenado el color junto con la información acerca del espaciado.

¿Cómo se produce el efecto McCollough? Nadie lo sabe con detalle, pero se ha sugerido algún que otro modelo elemental. En la figura 4 se ilustra uno de ellos, propuesto por un grupo de estudiosos del sistema visual. Es un modelo rudo, en el sentido de que se desconocen los detalles de cómo funciona. A la izquierda hay una sección de la retina iluminada por un retículo horizontal de rayas negras y verdes. Los circuitos de la derecha representan las primeras etapas de la visión. Las señales se desplazan hacia la derecha para alcanzar los niveles más altos del proceso cerebral. Téngase en cuenta que la ilustración es sólo representativa. Ignoro dónde se encuentran las secciones analizadoras del sistema visual, si es que realmente pueden localizarse. Tampoco conozco con detalle su manera de funcionar.

Las señales relativas al color de las líneas verdes se inician en los fotorreceptores cónicos de la retina. Estos conos son de tres tipos, sensible cada uno a una parte del espectro visible. Una vez excitados los conos, se comparan sus señales, lo que resulta en la transmisión de otras señales hacia el interior del sistema visual. Estas nuevas señales indican la luminosidad y las intensidades relativas del rojo respecto al verde y del amarillo respecto al azul. En alguna etapa posterior, posiblemente interior al cerebro, las nuevas señales se comparan en función de los colores complementarios en pugna.

Por ejemplo, el verde y el magenta compiten en esta etapa posterior del análisis. Consideremos esta competencia en términos de números asignados a las intensidades de las señales: positivos al verde, negativos al magenta y cero para la mezcla a partes iguales. Cuando el ojo ve verde, la intensidad de la señal podría ser, pongamos por caso, de +200. Cuando ve magenta, la intensidad podría ser de −200. Si se detectan cantidades iguales de ambos colores, la intensidad de la señal cromática será cero.

Los bordes de las rayas los detectan grupos de fotorreceptores que yacen en el lugar de la retina, o muy cerca de ella, donde se produce la imagen del borde. La información procedente de un gran número de detectores del borde alimenta un detector de retículos, el cual envía hacia el interior del sistema visual una señal de que se ve un retículo de una cierta orientación y una cierta periodicidad. Nadie sabe con certeza cómo actúa el detector de retículos. Supongamos que reconozca un retículo comparándolo con algunos retículos patrón. Cuando se detecte una identidad se enviará una señal por una línea de salida. La línea de salida elegida dependerá de cuál de los retículos patrón concuerde con el visto. Si éste es horizontal, con un determinado espaciado, se enviará una señal por una línea de salida. Si rotamos el retículo hasta ponerlo vertical, la señal se enviará por otra línea. Si el espaciado se varía lo suficiente para que el retículo concuerde con otro patrón, la señal saldrá por una tercera línea.

El efecto McCollough puede explicarse en términos de este modelo. Cuando se ve un retículo horizontal de rayas negras y verdes, se envían señales fuertes por dos líneas. Una de las señales atañe al verde de la luz y la otra a la orientación y espaciado del retículo. Entre las dos líneas comienza a establecerse una interacción inhibidora; ésta debilita la señal del verde. Inhibición que no llegamos siquiera a advertir bajo una luz ambiental normal. Pero aunque la intensidad de esta interacción se fortalezca lo suficiente en sólo unos cinco minutos, puede persistir durante horas e incluso días.

Supongamos que se mira después a un retículo sin color de igual orientación y espaciado que el primero. En la línea de color la señal es cero porque el retículo tiene rayas blancas. Si la señal cromática continuase siendo cero, la percepción sería blanco. Pero la inhibición que impone la interacción hace que la intensidad de la señal sea negativa. Y así, la señal cromática registra magenta, que será el color que se perciba en el retículo sin color.

En la línea de transmisión del color la inhibición se impone únicamente si el detector de retículos acciona la línea de salida cuya interacción está intensificada. Si se acciona alguna otra línea de salida, la de transmisión de color no se altera. Supóngase que se gira el retículo hasta colocarlo vertical. Tras ello, la señal procedente del detector de retículos saldrá por otra línea, la línea de transmisión de color no se inhibirá y desaparecerá la coloración de McCollough.

Aunque este modelo se ajusta a numerosos experimentos relativos al efecto McCollough, es poco satisfactorio. ¿Por qué debe inhibirse la línea de color con la línea de salida de un detector de retículos? Me preocupa asimismo el siguiente experimento. Si el modelo fuese correcto, debería formarse la interacción inhibidora al trasladar la vista por encima de un retículo sin color colocado al lado de una zona verde. Pero de esta manera no se consigue producir la coloración.

Para divertirme un poco ideé un modelo distinto para el efecto McCollough [*véase la figura 5*]. Ignoro si alguien ya lo ha estudiado antes. Difiere del precedente en que el detector de retículos es de por sí sensible al color, y envía una señal cromática al analizador que compara los colores complementarios. El analizador suma la señal procedente del detector de retículos con la cromática principal procedente de la retina. La intensidad de la señal cromática procedente del retículo depende de que éste se haya reconocido.

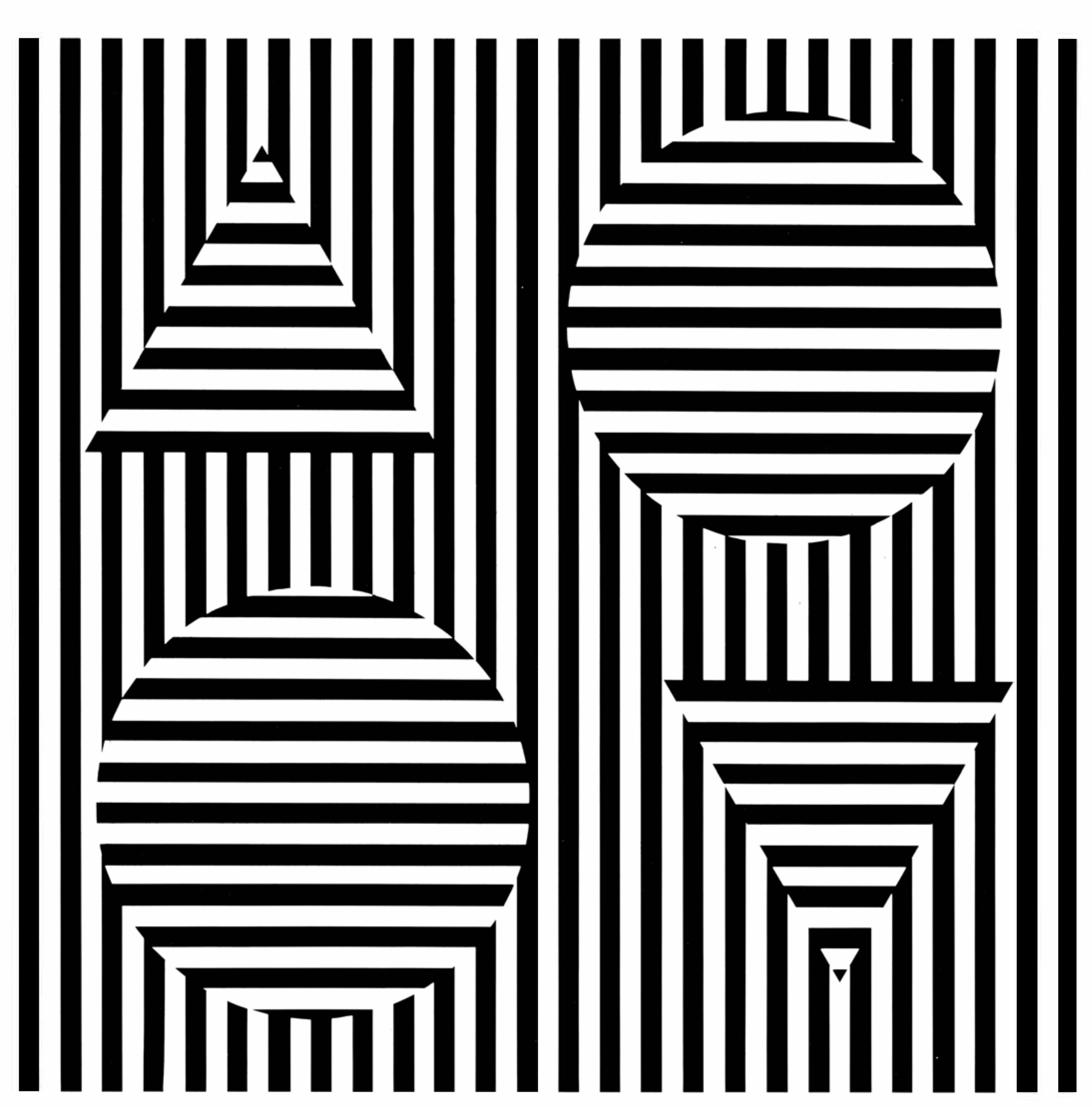
La sensibilidad al color del detector de retículos puede que sea una consecuencia ineludible del hecho de que

tanto él como los detectores de bordes analicen señales procedentes de células cromosensibles de la retina. La detección de un borde quizá dependa en parte del contraste entre los colores que haya a un lado y otro del mismo. Por ello, no debe sorprender una señal acerca del color procedente del detector de retículos.

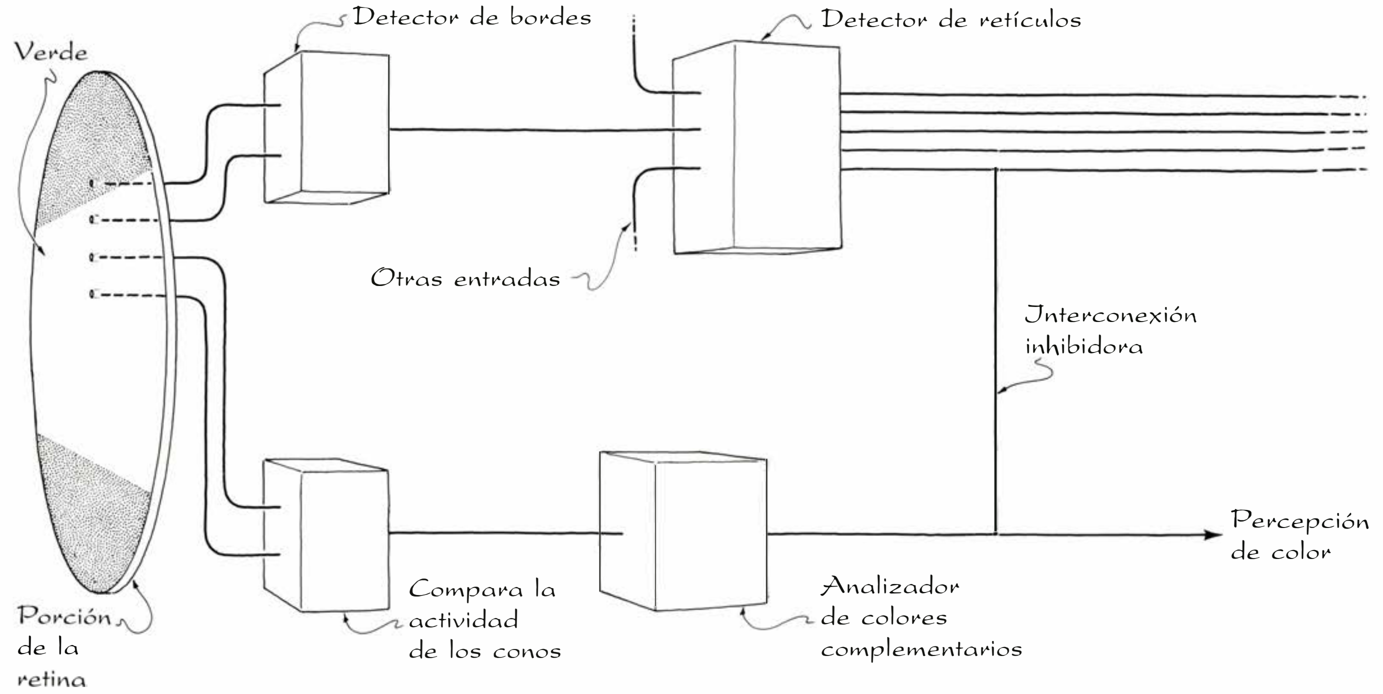
Supongamos que se mira un retículo horizontal de rayas negras y verdes. Tras identificarlo con un retículo patrón, el detector de retículos manda, por una línea de salida, una señal relativa a la orientación y periodicidad del retículo. A causa de esa identificación, el detector envía al analizador cromático una señal acerca del verde presente en el retículo. Supongamos que la señal directa sea de +200 unidades y que la señal verde procedente del detector de retículos sea de +10. La señal que surja del analizador cromático será de +210 unidades y se percibirá como verde.

Según se prosigue el examen del retículo, el detector reacciona menos tanto al retículo como a su color. Un rato después, el detector comienza a reducir su señal relativa a las características del retículo y la percepción de éste se hace más deficiente. Además, el detector empieza a inhibir la señal del verde, enviando en su lugar una señal magenta al analizador cromático. Supongamos que la intensidad de esta señal sea de −10 unidades. Si la señal directa del verde vuelve a ser de +200, la señal que surja del analizador cromático será de +190, o sea, un verde algo más pálido que el anterior.

Cuando se sustituye el retículo coloreado por otro sin color de igual espaciado y orientación, el detector de retículos lo reconocerá y volverá a en-



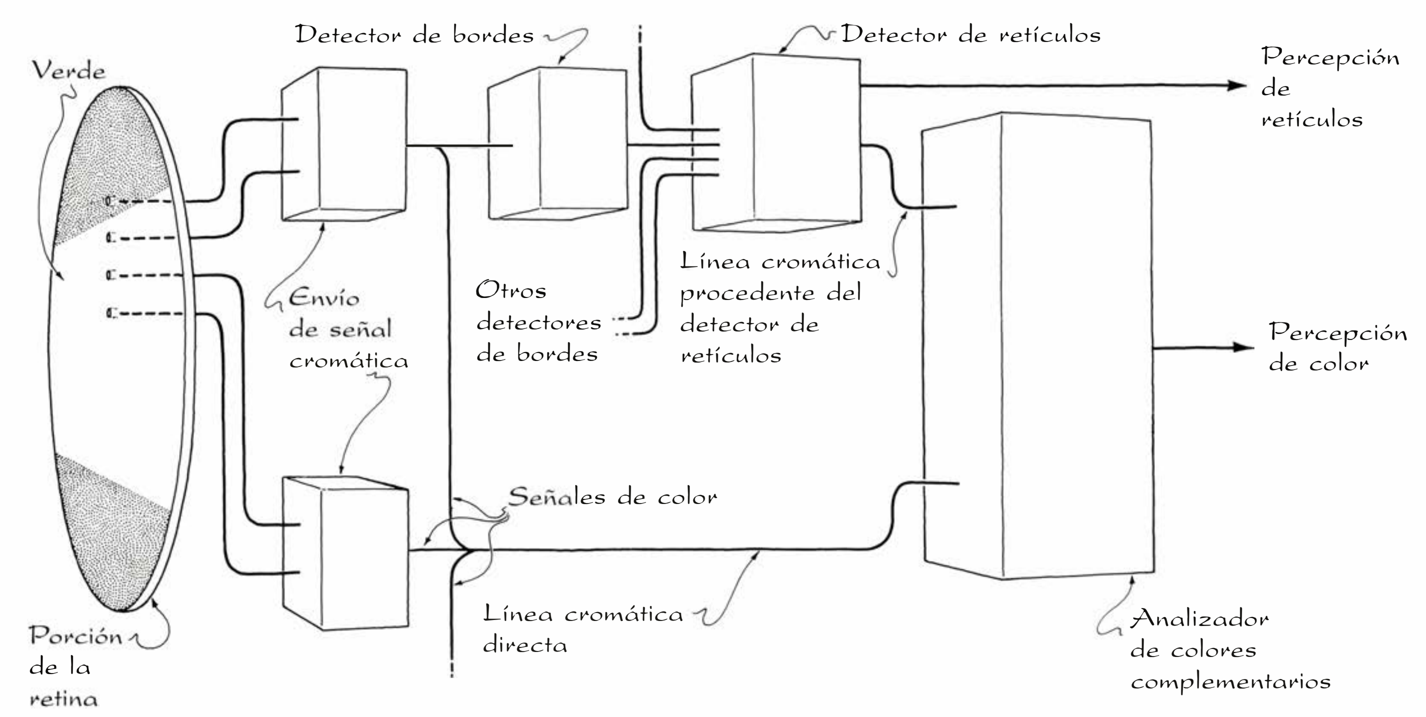
*3. Retículos que producen el efecto McCollough*




*4. Modelo de inhibición de una señal cromática*

viar, debido a su inhibición al verde, una señal de magenta al analizador cromático. Adviértase aquí que el color procedente del detector de retículos se activa al reconocer la presencia de un retículo, no mediante la intensidad de señal cromática alguna que proceda de los detectores de bordes. Si la intensidad de la señal magenta fuese de −10 y la correspondiente a la señal cromática directa cero (en virtud del color blanco de la raya observada), el analizador cromático remitiría una señal de −10: un magenta tenue. De aquí que al contemplar el retículo sin color se perciba un magenta pálido superpuesto a las rayas blancas. Esta coloración es el efecto McCollough.

¿Qué pasa al observar el retículo sin color bajo luz débil? Pues que la coloración se acentúa más, no porque la señal sea más intensa, sino porque la iluminación de las rayas blancas es más pobre, y por ello se distingue mejor el colorido magenta. Esta coloración sigue manifestándose aun cuando la luz sea tan tenue que los conos ya no actúen. Con esta luz, se hacen cargo de la visión las células del aparato visual llamadas bastoncitos. Estas no envían información cromática, pero sí sirven para detectar bordes y así emitir una señal al detector de retículos. Y como éste sigue fatigado, debido a la anterior observación de un retículo, la señal que envía al analizador cromático sigue siendo de −10 unidades de magenta. Así continúa percibiéndose la coloración.




*5. Modelo con detector de retículos sensible al color*

Imaginé, siguiendo este modelo, que bajo unas condiciones de iluminación adecuadas una señal de magenta procedente de un detector de retículos fatigado podría anular una pequeña señal directa de verde. Por ejemplo, si la señal magenta es de −10 y la señal directa de verde es de +10, su suma en el analizador cromático sería cero. En tal caso, las rayas verdes del retículo aparecerían grises, o sea, sin color y pálidas. ¿Tiene lugar realmente ese proceso de adquisición grisácea? Si es así, ¿reaparecería el verde al rotar el retículo de modo que las rayas estuvieran verticales, viéndose el detector de retículos obligado a efectuar un nuevo proceso de reconocimiento? Supongamos que haya una mancha verde aislada en otro lugar del campo visual. ¿Seguiría viéndose verde aunque se torne gris el verde del retículo?

Preparé el experimento añadiendo algunas manchas a un lado de mi habitual retículo verde y negro. De las manchas, unas eran negras y otras verdes; unas pequeñas y, otras, franjas del mismo largo y ancho que las del retículo. Atenué las luces de la habitación, esperé 15 minutos para que los ojos se me adaptaran a la iluminación y examiné entonces el retículo durante 20 minutos.

Aunque inicialmente pude distinguir el verde en el retículo, el color no tardó en diluirse en gris. El propio retículo se hizo más difícil de percibir. En aquel momento, el detector de retículos de mi cerebro debió hacerse menos reactivo y su sensibilidad al color mutarse a una inhibición al verde. Con todo, me pregunté si el gris resultaría de una disminución de la señal directa del verde, porque los conos se estuvieran cansando del color. Comprobé esa posibilidad llevando la vista a las manchas verdes aisladas. Este verde seguía siendo perceptible. Lo comprobé también rotando el retículo hasta poner las rayas verticales: inmediatamente, reaparecieron las rayas verdes. Por estas razones, el gris de las rayas en la orientación inicial parece deberse a la inhibición del verde dentro del detector de retículos.

En artículos ya publicados se describen muchos más experimentos en torno al efecto McCollough. Quizá alguien pueda idear alguno original y construir un modelo mejor que los considerados aquí. Si es así, me agradará recibir noticias suyas. Me interesarían, especialmente, experimentos que refuten mi modelo del detector de retículos sensible al color.

# Libros

## *Partículas subatómicas, monocotiledóneas, economía española y Stephen Hawking*

Luis Bru, Oriol de Bolós, Antonio Domínguez y Jesús Mosterín

PARTÍCULAS SUBATÓMICAS, por Steven Weinberg. Prensa Científica, S.A. Barcelona, 1985. Divulgar la ciencia, la física en particular, no es tarea fácil. Para conseguirlo se precisa que concurran en el autor una serie de circunstancias poco corrientes: un conocimiento muy profundo de la materia, poseer una pluma fácil y olvidarse del prestigio que pueda poseerse y saber descender al nivel de conocimientos de los posibles lectores, que suele abarcar un campo muy extenso. Desde los auténticos profanos, que desean ensanchar su nivel cultural, dotados de una muy envidiable curiosidad, hasta otros, auténticos especialistas, que suelen gozar con la lectura de esa literatura. También están los intermedios, que son, con toda seguridad los más, que conocen algo del tema, pero que no han podido penetrar en él por falta material de tiempo o por no poseer la base matemática que precisarían.

En el campo de la física ha habido en este siglo un plantel de magníficos divulgadores. Citemos, por ejemplo, a Eddington, Gamow o Sir James Jeans, de un lado y a Russell de otro. En España no puede olvidarse a Miguel Catalán.

Pues bien, aparece ahora otro auténtico maestro en el arte de divulgar la ciencia: el norteamericano Steven Weinberg, de la Universidad de Texas, que compartió el premio Nobel de física de 1979 con Sheldon Glashow y A. Salam, por su "contribución a la teoría de la unificación de las fuerzas débiles de interacción y electromagnética entre las partículas elementales". Es, además, un experto en historia medieval y el único físico de la Academia Americana Medieval.

El libro ha sido traducido por dos físicos, Ana Gómez Antón y J. M. Sánchez Ron, quienes, estoy seguro, habrán pasado ratos deliciosos mientras llevaban a cabo su tarea, modélica en todos sus aspectos, y a los que felicito por la pulcritud y acierto con que se ha realizado.

El mundo de las partículas elementales puede calificarse de fabuloso. He tenido la suerte de haber seguido su desarrollo a lo largo de mi dilatada vida científica y asistido al crecimiento de una familia, reducidísima al principio, convertida en la actualidad en numerosa, compuesta por docenas y docenas de miembros. Hasta tal punto ha aumentado el número de sus integrantes, que ha habido que agruparlas en subfamilias, cada una de las cuales constituida por individuos que se comportan de una manera específica. De forma precisa, elegante e incluso con "un cierto suspense", lo condensa Weinberg en su magnífico libro.

Quizás el mayor acierto del autor radica en haber escrito un libro que, en realidad, se compone de dos distintos, inteligentemente entretejidos. Uno recoge la secuencia de los hitos de la física del siglo XX que indujeron al descubrimiento de las partículas subatómicas: el *electrón*, el *protón* y el *neutrón*. El otro constituye una introducción a las parcelas de la física clásica (mecánica, electricidad y magnetismo, calor) que intervienen en estos descubrimientos. Las nociones fundamentales se van incorporando allí donde resultan necesarias para la comprensión del relato histórico, apoyándose en las descritas con anterioridad. Con Weinberg el lector revivirá la sucesión histórica de experimentos que disecaron el átomo, compartiendo así la empresa intelectual que ha remodelado la física del siglo XX. El autor ofrece una introducción que distingue este libro de la mayoría de los publicados hasta ahora, permitiendo a los lectores profanos en la materia de adquirir una base cultural de los revolucionarios conceptos y descubrimientos científicos de nuestros días. También el profesional de la física se sentirá fascinado, y así me ha sucedido a mí, con la exposición de los hallazgos cruciales que se describen en la obra, pues constituyen una parte del sustrato sobre el que se levantaron los suyos propios más recientes.

Weinberg evoca una imagen vívida de los actores que en ella intervinieron, nombres que siguen citándose cada día en las revistas especializadas así como en los libros más recientes. Sitúa gran parte de la acción en un escenario, el laboratorio Cavendish de la Universidad de Cambridge. De él dice textualmente: "Fue el marco, no tanto de un asalto general de los secretos de la naturaleza, cuanto de una guerra de guerrillas, de un esfuerzo dotado de pocos recursos, en el cual las armas principales eran la inteligencia y el ánimo de seres especialmente dotados".

El libro, además de un prefacio cuya lectura no tiene desperdicio, lo componen cinco capítulos que llevan los siguientes títulos: "Un mundo de partículas", "El descubrimiento del electrón", "La escala atómica", "El núcleo" y "Más partículas".

Cuenta con una gran cantidad de fotografías y gráficos originales, que hacen la delicia de los profesores rigurosos y enamorados de la historia de la física, destacando en doble plana los componentes del laboratorio Cavendish en 1933, rodeando a esos dos colosos que fueron Sir J. J. Thomson y Lord Rutherford. De ellos, un número elevado consiguieron en su día ser galardonados con el premio Nobel.

No voy a pormenorizar el contenido de cada uno de los capítulos del libro de Weinberg. Expondré al lector lo que considero más saliente del mundo de las partículas que componen lo que llamamos materia y el de aquellas otras responsables de "transmitir" las fuerzas que aceptamos que existen en la naturaleza, y que de manera clara y siguiendo cierto orden aparecen en el libro que estamos glosando.

La afirmación explícita de que la materia está compuesta de partículas llamadas átomos (que en griego significa no divisible) se debe a Demócrito de Abdera y a Leucipo en el siglo V antes de Jesucristo. De Abdera quedan sus ruinas. Ninguna de las palabras escritas por Leucipo han llegado hasta nosotros y de los alegatos de Demócrito sólo poseemos unos pocos fragmentos que no han prestado gran ayuda, pero la idea de atomismo, pese a sus muchos detractores, imperó durante veinticinco siglos. Con el advenimiento de la ciencia moderna, la idea de átomo se convirtió en soporte de teorías cuantitativas de la materia. A finales del siglo XIX, la idea del átomo le era familiar a la mayoría de los científicos, pero no aceptada de manera universal. Figuras de la ciencia alemana como Mach,

Boltzmann y Ostwald, entre otros, se resistían a incluirla en su filosofía.

El panorama sufrió un cambio drástico en las primeras décadas del siglo actual. La materia está formada por átomos constituidos por un núcleo central, que contiene casi toda la masa del mismo, y un conjunto de cargas eléctricas negativas –electrones– en número igual al de cargas positivas contenidas en el núcleo, que giran alrededor de éste. El esquema resultaba demasiado sencillo y efectivamente era así. Al descubrir Chadwick en 1932, en el famoso Cavendish, una nueva partícula neutra, llamada *neutrón*, cuya masa es prácticamente la misma del protón, el panorama cambió sustancialmente. En un principio se pensó que el neutrón estaba formado por un protón y un electrón, para poder explicar así la llamada radiación beta, formada sólo por electrones, pero Fermi descubrió al año siguiente la falsedad de tal asunción. El neutrón es una partícula fundamental y el proceso que tiene lugar en la emisión beta es análogo al que sucede en la emisión de luz por un átomo excitado. *Ni la partícula beta ni el fotón están presentes en el átomo hasta el momento en que son emitidas*. Hay que hacer resaltar que el descubrimiento de Chadwick, la teoría de Fermi y los aceleradores de partículas construidos por Cockroft y Walton en el Reino Unido y por Lawrence en USA, marcan el inicio de la era nuclear, que no la trata Weinberg en su libro, pues no tiene cabida en él. Sin embargo, los capítulos 4 y 5 contienen suficiente información sobre ella muy bien construida.

Ahora bien, con velocidad vertiginosa, el número de partículas, con sus correspondientes características, ha aumentado de manera alarmante, sucediendo además que aquellas que se creían elementales no lo son. El propio protón encierra otras subpartículas, llamadas *quarks*, que se presentan con seis variedades, pero dos de entre ellas son suficientes para formar un protón o neutrón.

Para que las partículas existan y la materia no desaparezca se requiere la presencia de fuerzas capaces de mantenerlas unidas. Los físicos tratan de explicar el origen y la evolución del universo de la forma más simplificada posible, tendiendo a considerar que existen dos únicas partículas, los mencionados quarks y los *leptones*. Asimismo, quieren unificar las cuatro fuerzas presentes en el universo en una sola. Es el ansiado empeño de la unificación de esas cuatro fuerzas: la *electromagnética*, una de cuyas manifestaciones es la luz; la *débil*, responsable de los procesos radiactivos; la *fuerte*, que mantiene unidas las partículas que integran los núcleos, y la más antigua conocida, la *gravedad*. Los quarks se comportan como los bloques que componen todas las piezas atómicas gobernadas por las fuerzas fuertes; por ejemplo, los protones. Los leptones son partículas atómicas que obedecen a aquellas fuerzas y carecen de componentes internos. La muestra más clara la constituye el electrón.

En lo concerniente a las fuerzas, ¿cuáles son los agentes encargados de trasmitirlas? Se sabe que son partículas intermediarias las que cumplen ese cometido. La responsable de la fuerza electromagnética es el *fotón*. Las asociadas a la fuerza débil son los *bosones* $W^+$, $W^-$ y $Z^0$. Las que transportan la interacción fuerte entre los quarks son los *gluones*. Sólo la partícula ligada a la gravitación, el posible *gravitón*, aún no ha sido puesto en evidencia, pero hay que tener en cuenta que esta fuerza no obedece a la mecánica cuántica. Pero si esta unión, tantas veces imaginada, fuera una realidad, se podría establecer una teoría cuántica de la gravitación que daría como resultado el descubrimiento del hipotético gravitón.

El deseo de la unificación de las cuatro fuerzas tropieza con la dificultad de que habría que reconciliar las dos grandes teorías que gobiernan el mundo de la física, la relatividad de Einstein y la mecánica cuántica. Ambas coexisten pacíficamente pero no han sido capaces de entenderse; ambas se reparten nuestra concepción del mundo, reinando cada una en su dominio, pero los físicos jamás han abandonado la posibilidad de hacerlas cohabitar con el objeto de concebir un universo coherente obedeciendo a un único principio unificador. Podemos decir que casi ahora mismo se está trabajando con ansia febril sobre la teoría de las *partículas cuerdas*. En ella el último constituyente de la materia se presenta bajo la forma de una cuerda (*string*) con la que tal vez sea posible conseguir la ambición suprema de la ciencia, reunir toda la materia y todas las fuerzas en una única visión, una *teoría del todo* ("theory of everything"), que suministraría una sola explicación para el conjunto de los fenómenos de la naturaleza, desde la gran explosión inicial hasta las galaxias, de la corriente eléctrica a la energía nuclear, de los agujeros negros a la radiactividad. El estudio que se está llevando a cabo podría poner en peligro la teoría general de la relatividad en lo referente a la identidad de masa inerte y masa gravitatoria y a la necesidad de aceptar un espacio de diez dimensiones.

En los capítulos 4 y 5 se resume que las partículas elementales que constituyen la materia se agrupan, en número de doce, en las dos familias ya mencionadas, los quarks y los leptones, aun cuando se sigue pensando que ellas mismas podrían estar compuestas por partículas todavía más pequeñas.

En el libro aparecen numerosos ejemplos numéricos, imprescindibles para los lectores profanos, que pueden adquirir así un concepto tan importante como es el de orden de magnitud. También algunas disquisiciones de un carácter matemático elemental y un apéndice de valor inestimable para todos los que cuentan con alguna preparación en el campo de la física.

Las palabras finales de Weinberg encierran una gran belleza unida a una profunda realidad. "No me gustaría –dice– que el lector se quedara con la impresión de que la rama de la física de las partículas ha degenerado en una suerte de coleccionismo de mariposas, con la particularidad de que las nuestras no viven lo suficiente para cazarlas en la naturaleza y hemos de crearlas en el laboratorio del entomólogo. Sería un juicio erróneo. Una vez que se resolvió la vieja pregunta con el descubrimiento del electrón, el protón y el neutrón, el problema mudó su planteamiento. El verdadero propósito que guía nuestros estudios experimentales y teóricos sobre las partículas elementales no es confeccionar una lista exhaustiva de ellas y sus propiedades. Antes bien, nos guía el afán de comprender los principios subyacentes que dictan por qué la naturaleza –partículas, núcleos, átomos, rocas y estrellas– es como es. Nuestra experiencia demuestra que el estudio de las partículas elementales constituye hoy el mejor modo, tal vez el único, de dar con las leyes fundamentales que rigen el universo. Cuando el poeta William Blake necesitó resumir la ciencia en una línea habló de los átomos de Demócrito y de las partículas de luz de Newton. De la época de Demócrito y de Leucipo a la de Blake y a la nuestra, la idea de la partícula fundamental ha sido siempre emblema de la meta más ansiada de la ciencia: comprender la complejidad de la naturaleza en términos sencillos."

En esos mismos capítulos el autor introduce al lector en el mundo de las

antipartículas, que están desempeñando un papel decisivo en la mejor comprensión de muchos hechos y fenómenos, en el de la inestabilidad y cortísima vida de algunas partículas, así como la idea, en vía de confirmación, de que el protón, ejemplo de partícula estable, podría desintegrarse en un plazo del orden de $10^{30}$ años, lo que conlleva la siguiente pregunta: ¿Será así, con la muerte de los protones, el fin descorazonador de toda la inmensidad del universo, del que sólo contemplamos el 10 por ciento de su totalidad? La muerte se asocia siempre con el frío, pero ¿podrá terminar todo tan friamente como la muerte de un simple protón? Esperemos que no sea así; mantengamos la esperanza de que aún quede algo de calor, cuando menos el que el espíritu es capaz de aportar. Aguardamos con impaciencia que todo esto y mucho más sea expuesto con la misma elegancia y sobriedad que dominan en este libro en una segunda parte anunciada por el propio Weinberg. (L.B.)

THE FAMILIES OF THE MONOCOTYLEDONS. STRUCTURE, EVOLUTION AND TAXONOMY, por R. M. T. Dahlgren, H. T. Clifford y P. F. Yeo, en cooperación con R. B. Faden, N. Jacobsen, K. Jakobsen, S. R. Jensen, B. J. Nielsen y F. N. Rasmussen. Springer-Verlag; Berlín, 1985. La gran elevación del nivel de conocimientos acerca de la constitución de los seres vivos que se ha producido en los últimos tiempos debe traducirse, necesariamente, en una remodelación y adaptación de los sistemas taxonómicos; en el caso de las plantas superiores, éstos solían apoyarse tradicionalmente en la consideración de un pequeño número de caracteres, correspondientes ante todo a la morfología de la flor y del fruto.

La artificialidad de la división simplista de las liliifloras, basada en primer lugar en el número de estambres y en la posición del ovario, ya fue reconocida, por ejemplo, hace muchos años. Así, en 1933, S. D. McKelvey y K. Sax, al considerar el número y la configuración de los cromosomas, pudieron demostrar la proximidad filogénica de *Yucca* y *Agave*, vegetales que, a pesar de mostrar afinidades evidentes en el porte, en la existencia de engrosamiento secundario, etc., se situaban en las liliáceas y en las amarilidáceas respectivamente, por poseer ovario súpero en el primer caso y ovario ínfero en el segundo.

La importante obra de que tratamos intenta poner al día la sistemática de las monocotiledóneas utilizando para ello todos los elementos de juicio disponibles hasta el momento. Con este fin se aplican los criterios y métodos de la escuela cladística, que fue definida en 1950 por W. Hennig, aunque ya tenía precedentes claros en la obra de R. von Wettstein y de muchos otros taxonomistas modernos. Esta escuela pretende construir el sistema mediante criterios exclusivamente filogenéticos.

En la obra que comentamos los autores marcan nítidamente las distancias respecto a los sistemas de clasificación sintéticos, que en el texto reciben también el nombre de eclécticos, los cuales, ante la dificultad de aplicar los criterios filogenéticos, que en muchos casos prácticos no pueden conducir a conclusiones seguras, intentan combinar dichos criterios con otros de carácter fenético, es decir, basados en semejanzas que no se sabe que expresen proximidad filogenética. Recuérdese que la posición sintética es sostenida, por ejemplo, por F. Ehrendorfer en el *Tratado de Botánica* de Strasburger e inspira los trabajos de Emberger y de otros autores europeos.

En un primer capítulo del libro que analizamos (págs. 1-16) se definen los conceptos fundamentales, tanto de morfología externa, como de anatomía y de funcionalismo, que se van a emplear.

De acuerdo con la tradición de la lengua inglesa, los autores siguen empleando el término *ovule*, biológicamente poco correcto. El tratamiento de las inflorescencias difiere considerablemente del tradicional. Las *inflorescencias determinadas* (o cerradas) pueden ser *panículas*, en las que las flores laterales crecen a distintos niveles por debajo de la flor apical, y *cimas*, en las que las flores laterales nacen en un nivel único por debajo de la flor apical. Las *inflorescencias indeterminadas* (o abiertas) se dividen en *tirsos*, en los que los pedicelos de las flores laterales originan nuevas flores en la axila de bractéolas, e *inflorescencias racemosas* (o *botrios*), en las que ello no sucede; entre estas últimas se encuentran el *racimo*, la *umbela*, la *espiga*, el *espádice* y el *capítulo*.

Un capítulo muy claro e ilustrativo (págs. 17-22), redactado por S. R. Jensen y B. J. Nielsen, se refiere a los caracteres químicos. En él se pasa revista, de un modo bastante general, a las posibilidades y a las limitaciones de la llamada quimiotaxonomía.

Las págs. 23-43 se dedican a una exposición de los criterios básicos de la cladística y de los conceptos generales sobre evolución aplicables a la taxonomía. Se tratan en detalle, como ejemplo, algunos de los casos críticos que se presentan dentro de las monocotiledóneas, cual los de las lemnáceas, de la inflorescencia y la flor de *Carex*, de las relaciones entre ninfeales y alismatales, etcétera.

Las cuatro páginas siguientes abordan los caracteres generales de las monocotiledóneas. Se admite que constituyen un grupo monofilético, caracterizado en primer lugar por dos caracteres: presencia de un solo cotiledón y existencia de cuerpos proteínicos triangulares en los plastos de los tubos cribosos. A ellos se añaden los demás caracteres que se indican en los manuales de botánica y que aquí se examinan en detalle y de modo comparativo.

El origen de las monocotiledóneas se estudia en las págs. 48-57. Se descarta la hipótesis de W. C. Burger (1981), según la cual las primeras angiospermas, derivadas de pteridospermas o de cicadófitos, serían vegetales próximos a monocotiledóneas, a partir de los que se habrían originado las dicotiledóneas.

Los autores se inclinan por la hipótesis contraria, generalmente admitida por los botánicos actuales, según la cual las monocotiledóneas procederían de dicotiledóneas primitivas relacionadas con el grupo de las magnoliifloras-ninfeifloras. Pero consideran demostrado (págs. 42-43) que las llamativas semejanzas que existen entre ninfeales y alismatales responden ante todo a convergencia y que la hipótesis de un parentesco próximo entre los dos grupos no resiste un análisis crítico. Las dioscoreales, en las que los autores incluyen *Smilax* y *Paris*, parece que son el grupo que mayores afinidades presenta respecto a las dictotiledóneas primitivas (magnoliifloras).

En la página 58 se exponen brevemente las monocotiledóneas fósiles, y se reconoce que no son de gran ayuda para el conocimiento de la filogenia del grupo. Los primeros granos de polen con aspecto de monocotiledónea aparecen en el Albiense (Cretáceo inferior), hace unos 115 millones de años. Los primeros pólenes de angiosperma datan de una época no muy anterior (122 millones de años).

El capítulo siguiente (págs. 59-75) acomete el estudio de la distribución de los distintos caracteres dentro de las monocotiledóneas. A continuación (págs. 76-90) se hace una valoración de la significación ancestral o derivada de

los distintos estados de cada carácter. Las páginas 91-105 se destinan a estudiar la evolución dentro de las monocotiledóneas. Si se admite la hipótesis de que sean derivadas de dicotiledóneas primitivas, las dioscoreales (pertenecientes al superorden liliifloras) serían el grupo central de las monocotiledóneas, aunque sus representantes actuales no corresponden exactamente a los que se supone primitivos.

A partir de monocotiledóneas primitivas próximas a las dioscoreales, la evolución se habría producido en varias direcciones. Una rama habría dado origen al resto de las liliifloras. Una línea distinta lleva a las arifloras, triuridifloras y alismatifloras. Otra, a las bromeliifloras (en las que se incluye las tifales), zingiberifloras y commelinifloras (con commelinales, hidatelales, ciperales y poales). Quedan en posición problemática arecales, ciclantales y pandanales.

El resto del libro (págs. 106-482; las páginas finales, hasta 520, se destinan a bibliografía e índices) comprende la descripción detalla de los superórdenes, órdenes y familias de las monocotiledóneas. Se da noticia también de los géneros y especies principales. Acompañan claves dicotómicas para el reconocimiento de las familias.

El sistema presentado se aparta mucho de las concepciones tradicionales. El número de familias resulta extraordinariamente aumentado. Los mismos autores reconocen (pág. V) que muchos botánicos les van a considerar "splitters". La clásica familia de las liliáceas, por ejemplo, se descompone, no ya en familias distintas, sino en varios órdenes: las esmilacáceas, con *Smilax*, y las triliáceas, con *Trillium* y *Paris*, se incluyen en el orden de las dioscoreales; convalariáceas, asparagáceas, ruscáceas, dracenáceas, agaváceas, hemerocalidáceas, asfodeláceas, antericáceas, afilantáceas, hiacintáceas, aliáceas, amarilidáceas, etcétera, integran el orden asparagales; colquicáceas, liliáceas, iridáceas, etcétera, forman parte del orden liliales. Las orquídeas se descomponen en tres familias: apostasiáceas, cipripediáceas y orquidáceas.

La obra aporta un gran acopio de observaciones e interpretaciones, de las que no se podrá prescindir en el futuro. Su aparición significa un importante adelanto en el conocimiento de las monocotiledóneas. No tendría sentido intentar dar un juicio, necesariamente superficial, sobre detalles concretos de un libro que es resultado del trabajo intenso y prolongado de un equipo de especialistas eminentes. Unicamente querríamos dejar abierta la cuestión acerca de si hubiera sido posible expresar lo fundamental de las mismas interpretaciones filogenéticas con el empleo de unidades taxonómicas de rango inferior.

La obra no se detiene apenas en las especies. Parece, sin embargo, que el criterio pulverizador que prevalece en la delimitación de órdenes y familias no se extiende a las unidades taxonómicas de rango inferior. Así, por ejemplo, se dice que el género *Festuca* comprende 80 especies, mientras que la exposición sumamente pulverizadora de I. Markgraf-Dannenberg en *Flora Europaea* (vol. V, págs. 125-153, 1980) reconoce 170 solamente para el área estudiada. En la pág. 450 de la obra que reseñamos se habla también de *Festuca rubra* var. *rubra* y var. *commutata*, denominaciones que no suelen verse en los textos de los botánicos pulverizadores actuales.

Es cierto que el problema de la amplitud de los taxones se presenta de modo muy diferente en los distintos niveles. En lo que concierne a las unidades superiores, se trata simplemente de poner de manifiesto unas relaciones filogenéticas de parentesco más o menos próximo, y la adopción de un criterio u otro afecta sólo a los tratados de botánica y a las discusiones de los especialistas. La especie, en cambio, es una unidad de utilización práctica, empleada en el quehacer diario no sólo por los especialistas en taxonomía, sino por gran número de personas ajenas a la botánica. Por ello, aparte las consideraciones de orden científico (¿es realmente posible hacer coincidir sus límites con los de las poblaciones aisladas genéticamente por barreras intrínsecas, como pretenden muchos genetistas?), parece conveniente que su reconocimiento no presente obstáculos muy difíciles de superar. (O.B.)

MADRID Y LA ECONOMÍA ESPAÑOLA, 1560-1850, por David R. Ringrose. Alianza Editorial, Madrid, 1985. El autor de este libro es un hispanista norteamericano bien conocido por todos los que se interesan por nuestra historia económica. En 1970 publicó un libro que inmediatamente se convirtió en un clásico en la materia: *Transportation and Economic Stagnation in Spain, 1750-1850*. (Traducción española: *Los transportes y el estancamiento económico de España*, M. Tecnos, 1972.) Su tesis básica es que el desarrollo económico de España ha sido contrariado por un relieve muy pronunciado, que eleva enormemente el coste de los transportes terrestres a la vez que ha creado una compartimentación regional y comarcal que ha dificultado la formación de un mercado unificado. El llevó el análisis hasta la implantación de los ferrocarriles, pero es evidente que éstos, a pesar de su superioridad sobre las antiguas caravanas de mulos y carretas, no podían eliminar unos obstáculos que nacen de la propia geología peninsular.

Al mismo tiempo se preocupó por el fenómeno del crecimiento de una capital política, Madrid, que surge en medio de la Meseta como un desafío humano a los imperativos de la naturaleza: un oasis de prosperidad en un medio poco favorable; un centro único de decisión para un país con tendencias centrífugas; un sistema radial de comunicaciones que obliga a pasar las vías maestras por el centro desolado cuando la mayor vitalidad de las costas parecía exigir un sistema periférico. Para Ringrose ambos problemas, el de las dificultades estructurales del sistema de comunicaciones y el de la ubicación anormal de la capital, son inseparables y los ha estudiado conjuntamente. No es hombre que se haya desperdigado en variedad de temas; casi todos sus artículos y sus dos libros básicos se relacionan con éste, y por eso la obra cuya traducción comentamos aparece como la suma de todas sus investigaciones y reflexiones sobre el mismo.

*Madrid y la economía española* es un libro de muy denso contenido que se articula en tres partes; en la primera, titulada simplemente "Madrid", estudia la evolución de su población a lo largo de la Edad Moderna hasta mediados del pasado siglo, y la de su economía en el mismo lapso de tiempo. El estudio de la población se resiente de la falta de investigaciones sobre las fuentes primarias, que son los archivos parroquiales, aún numerosos a pesar de las destrucciones sufridas en los últimos tiempos. Sobre este punto esperamos que María Carbajo Isla nos ofrezca pronto el resultado definitivo de la investigación que viene llevando a cabo desde hace años. Ringrose nos presenta una serie de estimaciones de la población madrileña sacadas de diversos autores y fuentes, de valor muy desigual, y la crítica que hace de sus valores numéricos nos parece insuficiente, llegando a conclusiones no siempre aceptables; concretamente, el

salto desde los 65.000 habitantes que tendría Madrid al finalizar el siglo XVI hasta 175.000 en 1630 con posterior caída a 125.000 no se apoya en ninguna base documental y además es inverosímil. ¿Por qué esa pérdida de un tercio de la población cuando la capital estaba en plena fase de consolidación y crecimiento? Una caída tan brutal hubiera dejado huellas no sólo en las estadísticas, sino en la literatura de la época. Pues bien, esas huellas faltan por completo. A partir del siglo XVIII las fuentes resultan más fiables; se asiste en dicha centuria a un crecimiento moderado que situó la población madrileña en los umbrales de los 200.000 en 1800, seguido de un neto retroceso al que no fueron ajenos los avatares de la Guerra de la Independencia. El crecimiento recomenzó desde 1820, pero fue un crecimiento modesto en comparación con el de otras capitales europeas; los 300.000 habitantes del Madrid isabelino representaban poco frente a los dos millones del París del Segundo Imperio, y no digamos frente a los tres millones del Londres victoriano. Verdad es que Madrid nunca figuró entre las capitales europeas de primer rango; comenzó tarde y su carrera ascensional se quebró pronto.

La economía madrileña de aquellos siglos está estudiada con gran competencia y detalle utilizando, además de las fuentes clásicas, una amplia documentación, en especial la existente en el Archivo de la Villa.

La segunda parte del libro, bajo el rótulo "Madrid y la Castilla rural", es quizá la parte más original. En ella examina el impacto que el crecimiento de un núcleo urbano de gran capacidad adquisitiva tuvo en un amplio radio, atrayendo hacia sí los productos de las comarcas vecinas (e incluso de algunas bastante alejadas) ya por el libre juego del mercado, ya por los métodos compulsivos habituales en el Antiguo Régimen, como la obligación impuesta a numerosos pueblos de suministrar pan a la capital. La desorganización de las anteriores redes comerciales y la absorción de las fuerzas vitales de las ciudades vecinas provocó el dramático eclipse de ésta. Bien conocido es el caso de Toledo, rica y poblada en el siglo XVI, pobre y medio despoblada en el XVII; con más o menos intensidad, el caso se repitió en Valladolid, Avila, Segovia, Medina del Campo y otras muchas. Toda la Meseta sufrió las consecuencias de la creación de aquella capital de burócratas y rentistas que absorbía la sustancia del país sin dar nada a cambio. Casi su única industria exportadora era la editorial. La enorme diferencia entre lo que compraba y lo que vendía se saldaba con el importe de los sueldos de los funcionarios, las ganancias de los recaudadores y las rentas de los nobles y mayorazgos, que, a su vez, constituían una minoría dentro de la minoría, pues en el Madrid de mediados del XVIII el uno por ciento de los perceptores recibía el 40 por ciento de las rentas. Caso extremo de desigualdad dentro de una sociedad ya de por sí muy poco igualitaria.

La tercera parte ("Madrid y la economía española") generaliza las conclusiones anteriores a todo el espacio nacional; aunque no con tanta fuerza como su entorno inmediato, todo él sufrió el impacto negativo del centralismo madrileño. La conclusión de toda la obra puede cifrarse en el siguiente párrafo:

"Madrid fue para España una capital que era la imagen invertida del motor de desarrollo económico que Wrigley ve en el Londres de los siglos XVII y XVIII. Apenas proporcionó incentivos a la España Interior y funcionó más bien como 'motor' de su estancamiento económico. Concentró la riqueza, las rentas y la actividad comercial del país en un estrecho mercado constituido por unos pocos miles de familias acomodadas, de las que sólo eran realmente ricas unas escasas centenas, y por una muchedumbre de menesterosos con ingresos muy reducidos y una baja productividad" (páginas 133-134).

Este juicio negativo tiene ya sobrados precedentes en el pensamiento español, si bien entre nosotros se miró más bien desde el punto de vista político. Ringrose no entra a discutir si Felipe II hubiera hecho mejor elección situando la capital de sus estados en Lisboa, Sevilla o Barcelona, cuestión bastante ociosa. Se limita al aspecto económico y con un impresionante despligue de cifras y cuadros saca una conclusión negativa que puede suscribirse a condición de moderar su extremosidad. Es achaque común en los que se especializan en un tema magnificarlo, convertirlo en la llave maestra que desvela todos los secretos del pasado. Me temo que David Ringrose no haya sabido resistir a esta tentación. Es natural que el aumento de Madrid por motivos políticos perjudicara a otras de su entorno, singularmente aquellas que, como Toledo y Valladolid, eran aglomeraciones residenciales. También es natural (y positivo) que las *ferias* intermitentes de Medina del Campo hayan sido desplazadas por la actividad bancaria *permanente* de Madrid. Lo que no se explica es que la onda producida por el crecimiento de un nuevo centro comercial haya afectado gravemente a ciudades tan alejadas como Burgos, León o Albacete, y menos aún que la decadencia de la industria textil de Segovia y Cuenca tenga mucho que ver con el crecimiento de Madrid que, como el autor reconoce, no era un centro manufacturero. La decadencia mesetaria, su ruralización, tiene otras causas más profundas, estudiadas por autores como García Sanz, Pérez de Moreda y otros. Se hubiera producido igual sin la decisión tomada por Felipe II; incluso la situación de la España interior hubiera sido peor, pues, al fin y al cabo, la creación de un núcleo importante en el interior también tenía que ejercer efectos positivos.

Menos aún se explica que el crecimiento de un núcleo humano que no rebasó el 2 por 100 de la población española y el 3 por 100 de la renta nacional tuviera repercusiones tan serias sobre todo el conjunto económico español. Repito: en la tesis que defiende este libro hay mucha parte de verdad; es su exclusivismo lo que altera la visión de conjunto. No obstante, su impresionante acumulación de datos le confiere un interés y una autoridad tal que lo convierten en un clásico en la materia. Por encima de diferencias de interpretación hay que rendirse ante el colosal esfuerzo que ha realizado su autor y felicitarlo por la aportación que ha realizado a la historia de nuestro pasado. (A.D.O.)

**BEYOND THE BLACK HOLE. STEPHEN HAWKING'S UNIVERSE**, por John Boslough. Collins; London, 1985. Este librito (de 127 páginas), obra de un periodista científico americano, ofrece la primera exposición popular del pensamiento de uno de los físicos y cosmólogos más geniales de nuestro tiempo, Stephen Hawking.

Hawking nació en 1942, estudió física en Oxford y tras su licenciatura se trasladó a Cambridge. Para entonces había empezado a notar síntomas de una extraña enfermedad, que a veces le impedía atarse los zapatos o hablar con claridad. Diagnosticada en 1963 como esclerosis lateral amiotrófica (es decir, lenta y progresiva desintegración de las neuronas motoras, que regulan la actividad muscular voluntaria), los médicos pronosticaron a Hawking que sólo le quedaban dos años de vida. Deprimido, se encerró en su habitación en

Cambridge y, descuidando sus estudios, se dedicó a esperar la muerte escuchando música clásica, leyendo novelas de ciencia-ficción y bebiendo. Su tutor, el cosmólogo Sciama, consciente de la genialidad de su enfermizo pupilo, hacía la vista gorda y esperaba también.

Pasados los dos años previstos y en vista de que no se moría, superó su depresión, se casó y reanudó en serio sus investigaciones, que desde entonces no ha vuelto a interrumpir ni un solo día. Sin embargo, su condición física ha seguido deteriorándose. Sus músculos se han ido atrofiando. Desde 1970 vive confinado a una silla de ruedas. No puede mover brazos ni piernas. Ni siquiera controla del todo los movimientos de su boca. Su voz es apenas audible e inteligible, y necesita ser "traducida" por su secretaria o por algún discípulo muy acostumbrado a ella.

Este inválido, cuyo cuerpo está sometido a un deterioro implacable, posee una extraordinaria inteligencia. En su cerebro formula, manipula y retiene series larguísimas de complejas ecuaciones, y concibe ideas originales y audaces, que sus discípulos en Cambridge desarrollan y comprueban. Muchos lo consideran como el mayor físico teórico viviente. En su silla de ruedas ha ido a California y a Moscú, donde los cosmólogos escuchaban su voz apenas perceptible como si fuera un oráculo. En 1974 su silla de ruedas hubo de ser aupada por la escalinata de la *Royal Society*, de la que acaba de ser elegido miembro. Y en 1980 fue nombrado Profesor Lucasiano de Matemáticas de la Universidad de Cambridge, cátedra en la que le habían precedido Newton y Dirac.

Aunque las soluciones de Friedman a las ecuaciones de la relatividad general de Einstein exigían un universo dinámico e inestable, en expansión a partir de una singularidad o punto inicial, a principios de los años 60 esa idea seguía sin tomarse en serio. Entre 1965 y 1968 Hawking y Penrose probaron matemáticamente que (asumiendo la relatividad general) es necesario admitir una singularidad al principio del universo, un punto que concentraba toda la materia y cuya explosión (*big bang*) dio origen al universo, al espacio y al tiempo. El descubrimiento de la radiación cósmica de fondo y de la proporción de hidrógeno y helio en el universo vinieron a darles la razón. La teoría del estado estable cayó en desuso.

A continuación se dedicó Hawking a desarrollar la teoría de los agujeros negros, último estadio del colapso gravitatorio de una estrella suficientemente masiva. La velocidad de escape de un agujero negro es mayor que la de la luz, por lo que nada, ni siquiera la luz, puede escapar. La frontera que separa el mundo normal del abismo sin retorno del agujero negro se llama su horizonte de sucesos. En 1970 probó Hawking que el horizonte de sucesos de un agujero negro no puede decrecer. En 1971 postuló que el colapso de una estrella masiva no era la única manera como podía formarse un agujero negro. Durante el primer $10^{-20}$ segundo de "la gran explosión" podrían haberse formado millones de mini-agujeros negros, que persistían aún, invisibles, desde luego, pero contribuyendo a la masa del universo y, por tanto, a su destino final.

En 1973 Hawking combinó la relatividad general con la mecánica cuántica para probar que los agujeros negros tienen que emitir desde su horizonte un flujo constante de partículas, lo que luego ha pasado a llamarse la "radiación de Hawking". En el espacio pares de partículas y antipartículas se crean, se reúnen y se aniquilan, debido a efectos cuánticos. Pero en la zona del horizonte de un agujero negro una de las partículas es tragada por el agujero negro y la otra se queda sin aniquilar, escapando hacia afuera. Esta radiación de Hawking desempeña un papel fundamental en los modelos cosmológicos.

El autor del librito, Boslough, expone a un nivel muy elemental los desarrollos recientes de la física teórica y la cosmología, desde la unificación de las teorías de los diversos campos hasta la cuantificación de la relatividad general, comentándolos con Hawking y recogiendo sus observaciones y opiniones.

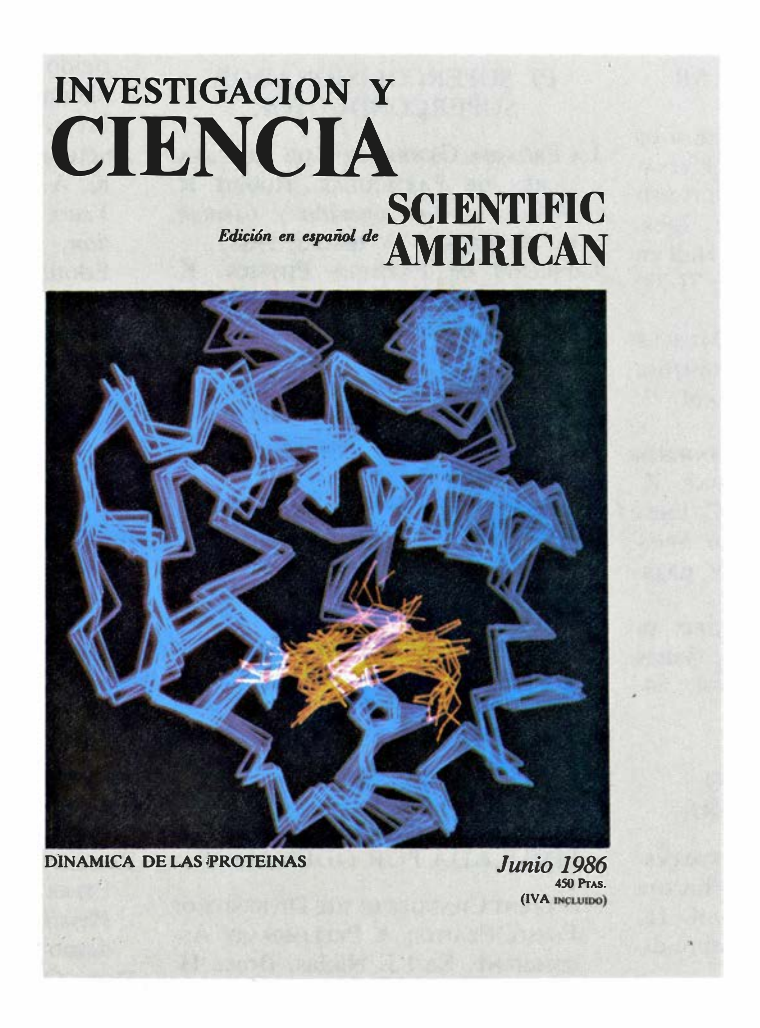
En los últimos años Hawking ha estado especialmente obsesionado con el tema del origen del universo, barajando en su cabeza las diversas teorías últimamente propuestas, incluidas las del modelo inflacionario o de burbujas. Precisamente en 1982 Hawking y Gibbons convocaron y reunieron en Cambridge la primera conferencia internacional sobre el segundo inicial del universo, a la que acudieron casi todos los cosmólogos activos del mundo. Hay bastante unanimidad a la hora de describir la evolución a partir del primer $10^{-33}$ segundo. Lo que queda (un $10^{-33}$ segundo) es prácticamente nada, y sin embargo todas las discusiones se centran en ese período inconcebiblemente minúsculo. "Me gustaría saber –dice Hawking a Boslough– qué pasó exactamente entre el $10^{-43}$ y el $10^{-33}$ primer segundo. Ahí se encuentra la respuesta última a todas las cuestiones sobre el universo." Y eso es lo que le interesa. "Mi meta es simple: lograr la comprensión completa del universo."

Su lección inaugural como Profesor Lucasiano, recogida como apéndice en este libro, se plantea la pregunta de si el final de la física teórica está ya a la vista. Y responde que sí. En su opinión faltan sólo unos 20 años para que logremos una teoría unificada y completa del universo, que dé cuenta de todas las partículas y todas las interacciones. La unificación de la interacción nuclear débil y del electromagnetismo en la teoría de Glashow, Weinberg y Salam es un primer paso, que ha encontrado rápido apoyo experimental con la reciente comprobación de los bosones $W^-$, $W^-$ y $Z^0$ en el CERN. Las teorías que unifiquen esta fuerza electrodébil con la fuerte, y luego con la gravedad, serán de casi imposible comprobación experimental directa, dadas las energías implicadas y dado el coste creciente de tales experimentos, consecuencia del éxito mismo de la física. Como dice Hawking: "Es un tributo a cuán lejos hemos llegado ya en la física teórica el que ahora sean necesarias máquinas enormes y una gran cantidad de dinero para llevar a cabo experimentos cuyos resultados no podamos predecir."

Hawking es optimista. La gran unificación la ve venir del lado de la teoría de supergravedad extendida $N=8$, y tendrá como resultado principal la cuantificación de la relatividad general y la superación de la actual incompatibilidad entre las dos teorías más potentes de la física: la relatividad general de los cosmólogos y la mecánica cuántica de los físicos de partículas. Todo lo cual no significa –advierte Hawking– que, el día que tengamos la teoría completa unificada, seamos capaces de hacer predicciones detalladas. Quizás sólo podamos hacer predicciones estadísticas sobre casos simples. Pero los secretos del cosmos y la materia habrán sido desentrañados. Y la física teórica ya no será tan emocionante.

Naturalmente no es nada seguro que lleguemos a una tal teoría unificada. ¿La necesita la física realmente? "Los físicos –comenta Boslough– necesitan una teoría unificada sólo en el sentido en el que Edmund Hillary necesitaba escalar el Everest." La *Biblioteca Científica Salvat* acaba de ofrecer la versión castellana de este ensayo. (J.M.)

# Seguiremos explorando los campos del conocimiento




**NUEVAS ESTRATEGIAS PARA LA ENERGIA NUCLEAR, por Richard K. Lester**
*Una posible vía para liberar a la energía nuclear del punto muerto en que se encuentra podría fundamentarse en una nueva generación de reactores, de baja potencia.*

**DINAMICA DE LAS PROTEINAS, por Martin Karplus y Andrew McCammon**
*Mediante simulaciones por ordenador se exploran los movimientos internos que subyacen a la actuación de esas moléculas, imprescindibles para el mantenimiento de la vida.*

**APLICACIONES DEL EFECTO HALL CUANTICO, por Bertrand I. Halperin**
*Permite ese efecto establecer el valor de ciertas constantes fundamentales con una precisión parangonable a la de las mediciones más exactas efectuadas hasta la fecha.*

**RITMOS INTERNOS EN LA EMIGRACION DE LAS AVES, por Eberhard Gwinner**
*Las aves migradoras disponen de un reloj interno que les dice cuándo empezar y cuándo acabar el vuelo. Además, les ayuda a dar con el punto de destino.*

**GIROSCOPIOS OPTICOS, por Dana Z. Anderson**
*Al igual que los de factura mecánica, detectan los movimientos de rotación, pero sin poseer apenas piezas móviles. Se emplean en los sistemas de guía en navegación.*

**LAS TENSIONES DEL NACIMIENTO, por Hugo Lagercrantz y Theodore A. Slotkin**
*El esfuerzo propio del recorrido del canal de nacimiento no perjudica a la mayoría de los recién nacidos. De hecho, la afluencia de hormonas "de tensión" favorece su superviviencia.*

**LA TROMPETA BARROCA, por Don Smithers, Klaus Wogram y John Bowsher**
*El análisis de la historia y las características físicas de ese olvidado instrumento permite que se interprete hoy como lo hicieron los músicos de los siglos XVII y XVIII.*

**COMUNICACIONES OPTICAS POR CABLES SUBMARINOS, por Raimundo Díaz de la Iglesia**
*Los cables de fibras ópticas transportan fotones y sustituyen con ventaja a los cables metálicos. Con ellos se entra en una nueva etapa del desarrollo de las comunicaciones submarinas.*



# Bibliografía

*Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:*


## UNION NEUROMUSCULAR

ASSOCIATION OF THE SYNAPTIC FROM OF ACETYLCHOLINESTERASE WITH EXTRACELLULAR MATRIX IN CULTURED MOUSE MUSCLE CELLS. N. C. Inestrosa, L. Silberstein y Z. W. Hall en *Cell*, volumen 29, páginas 71-79; 1982.

DIFFERENTIATION OF SKELETAL MUSCLE CELLS IN CULTURE. N. C. Inestrosa en *Cell Struct. & Funct.*, vol. 7, n.º 2, págs. 91-109; 1982.

SYNAPTIC BASAL LAMINA COMPONENTS MADE BY A MUSCLE CELL LINE. Z. W. Hall, L. Silberstein y N. C. Inestrosa en *Cold Spring Harbor Symposia on Quantitative Biology*, págs. 459-467; 1982.

ROLES OF EXTRACELLULAR MATRIX IN NEURAL DEVELOPMENT. J. R. Sanes en *Annu. Rev. Physiol.*, vol. 54, págs. 581-600; 1983.

## LA COLA DEL CAMPO MAGNETICO TERRESTRE

THE MAGNETOTAIL AND SUBSTORMS. C. T. Russell y R. L. McPherron en *Space Science Reviews*, vol. 15, n.º 2/3, págs. 205-266; noviembre-diciembre, 1983.

MAGNETIC RECONNECTION IN SPACE AND LABORATORY PLASMAS. Dirigido por Edward W. Hones, Jr. American Geophysical Union; 1984.

## GENETICA MOLECULAR DE LA HEMOFILIA

STRUCTURE OF HUMAN FACTOR VIII. Gordon A. Vehar, Bruce Keyt, Dan Eaton, Henry Rodriguez, Donogh P. O'Brien, Frances Rotblat, Herman Oppermann, Rodney Keck, William I. Wood, Richard N. Harkins, Edward G. D. Truddenham, Richard M. Law y Daniel J. Capon en *Nature*, vol. 312, n.º 5992, págs. 337-342; 22 de noviembre de 1984.

DETECTION AND SEQUENCE OF MUTATIONS IN THE FACTOR VIII GENE OF HAEMOPHILIACS. Jane Gitschier, William I. Wood, Edward G. D. Truddenham, Marc A. Shuman, Therese M. Goralka, Ellson Y. Chen y Richard M. Lawn en *Nature*, vol. 315, n.º 6018, págs. 427-430; 30 de mayo de 1985.

## EL SUPERCOLISIONADOR SUPERCONDUCTOR

LA PRÓXIMA GENERACIÓN DE ACELERADORES DE PARTÍCULAS. Robert R. Wilson en *Investigación y Ciencia*, n.º 42, págs. 6-23; marzo, 1980.

CONCEPTS OF PARTICLE PHYSICS. K. Gottfried y V. Weisskopf. Oxford University Press, 1984.

SUPERCONDUCTING MAGNETS FOR PARTICLE ACCELERATORS. R. Palmer y A. Tollestrup en *Annual Review of Nuclear and Particle Science*, vol. 34, págs. 247-284; 1984.

SUPERCOLLIDER PHYSICS. E. Eichten, I. Hinchliffe, K. Lane y C. Quigg en *Reviews of Modern Physics*, vol. 56, n.º 4, págs. 579-707; octubre, 1984.

THE TEVATRON. H. T. Edwards en *Annual Review of Nuclear and Particle Science*, vol. 35, págs. 605-660; 1985.

PARTÍCULAS SUBATÓMICAS. Steven Weinberg. Prensa Científica; Barcelona, 1985.

## EVOLUCION VEGETAL SIMULADA POR ODENADOR

APPARENT CHANGES IN THE DIVERSITY OF FOSSIL PLANTS: A PRELIMINARY ASSESSMENT. Karl J. Niklas, Bruce H. Tiffney y Andrew H. Knoll en *Evolutionary Biology*, vol. 12, págs. 1-79; 1980.

PALEOBOTANY AND THE EVOLUTION OF PLANTS. Wilson N. Stewart. Cambridge University Press, 1983.

MECHANICAL AND PHOTOSYNTHETIC CONSTRAINTS ON THE EVOLUTION OF PLANT SHAPE. Karl J. Niklas y Vincent Kerchner en *Paleobiology*, vol. 10, n.º 1, págs. 79-101; Winter, 1984.

## IMAGENES MENTALES Y SISTEMA VISUAL

IMAGE AND MIND. Stephen Michael Kosslyn. Harvard University Press, 1980.

LEVELS OF EQUIVALENCE IN IMAGERY AND PERCEPTION. Ronald A. Finke en *Psychological Review*, vol. 87, n.º 2, págs. 113-132; marzo, 1980.

MENTAL IMAGERY AND THE THIRD DIMENSION. Steven Pinker en *Journal of Experimental Psychology: General*, vol. 109, n.º 3, págs. 354-371; septiembre, 1980.

## ALMACENAJE DE GRANO EN ASSIROS

A FRIEND IN NEED IS A FRIEND INDEED: SOCIAL STORAGE AND THE ORIGINS OF SOCIAL RANKING. P. Halstead y J. O'Shea en *Ranking, Resource, and Exchange: Aspects of the Archaeology of Early European Society*, dirigido por Colin Renfrew y Stephen Shennan. Cambridge University Press,1982.

ANCIENT MACEDONIA: THE CULTURES. K. A. Wardle en *Macedonia—4000 Years of Greek History and Civilisation*, dirigido por M. Sakellariou. Edotike Athenon, 1982.

LINEAR B TABLETS AND THE MYCENAEAN ECONOMY. J. T. Killen en *Linear B: A 1984 Survey*, dirigido por A. M. Davies y Y. Duhoux. University of Louvain, 1985.

## EL EFECTO HALL CUANTICO

NEW METHOD FOR HIGH-ACCURACY DETERMINATION OF THE FINE-STRUCTURE CONSTANT BASED ON QUANTIZED HALL RESISTANCE. Klaus von Klitzing, Gerhard Dorda y Michael Pepper en *Physical Review Letters*, vol. 45, págs. 494-497, 11 de agosto de 1980.

THE QUANTUM HALL EFFECT. Klaus von Klitzing en *Europhysics News*, vol. 13, cuaderno n.º 2; abril, 1982.

ZUR ENTDECKUNG DES QUANTEN-HALL-EFFEKTES. Gottfried Landwehr en *Physikalische Blätter*, vol. 41, cuaderno n.º 11, pág. 357; 1985.

DER QUANTEN-HALL-EFFEKT. Janos Hajdn y Bernhard Kramer en *Physikalische Blätter*, vol. 41, cuaderno n.º 12, pág. 401; 1985.

## JUEGOS DE ORDENADOR

ARTIFICIAL INTELLIGENCE. Patrick Henry Winston. Addison-Wesley Publishing Co., 1977.

THE MISMEASURE OF MAN. Stephen Jay Gould. W. W. Norton & Company, Inc., 1981.

KNOW YOUR OWN I.Q. H. J. Eysenck. Penguin Books, 1984.

## TALLER Y LABORATORIO

COLOR ADAPTATIONS OF EDGE-DETECTORS IN THE HUMAN VISUAL SYSTEM. Celeste McCollough en *Science*,vol. 149, n.º 3688, págs. 1115-1116; 3 de septiembre de 1965.

FORM-SPECIFIC COLOUR AFTER EFFECTS IN SCOTOPIC ILLUMINATION. Charles F. Stromeyer III en *Nature*, vol. 250, n.º 5463, págs. 266-268; 1974.